URBAN ENERGY INFORMATICS: IMPROVING THE USABILITY OF BUILDING ENREGY DATA FOR COMMUNITY ENERGY EFFICIENCY

A Dissertation
Presented to
The Academic Faculty

by

Abigail Leigh Francisco

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Civil and Environmental Engineering

Georgia Institute of Technology
August, 2020

URBAN ENERGY INFORMATICS: IMPROVING THE USABILITY OF BUILDING ENREGY DATA FOR COMMUNITY ENERGY EFFICIENCY

Approved by:

Dr. John E. Taylor, Advisor
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Iris Tien
School of School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Emily Grubert
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Daniel Matisoff
School of Public Policy
Georgia Institute of Technology

Dr. Eric Marks
School of Civil and Environmental Engineering
Georgia Institute of Technology

Date Approved: April 17, 2020

ACKNOWLEDGEMENTS

I would not have completed this dissertation, let alone began it, without the support of many others to whom I am deeply thankful. Importantly, I would first like to thank my research advisor, JT (Dr. John Taylor). Your expertise, continued assurances, and reflective advice shaped my research trajectory considerably and challenged me to think more broadly about the scientific process.

I was also supported by a wonderful network of people at Georgia Tech. My conversations with my dissertation committee, including Dr. Emily Grubert, Dr. Eric Marks, Dr. Dan Matisoff, and Dr. Iris Tien, provided critical guidance that helped solidify my research questions and the interpretation of my results. Georgia Tech's Facility Management team—especially Jessica Rose—provided essential support and spent significant time answering my questions on building operations across campus and engaging with the results of my work. This journey would have been much worse off without the companionship from those at the Network Dynamics Lab. Chel, Lei, Susie, Yuli, Dimitri, Praga, and Neda, thank you for everything—for being sounding boards, moral compasses, and friends.

I am indebted to those who initiated this process. Back during undergrad, Leidy and Freddy planted the seeds for pursuing research and the PhD path, and they have lent advice and enlightening conversations throughout. The mentors, friends, and co-workers I met at Southface—a fantastically unusual place to start my working life—instilled in me a deep curiosity and excitement about buildings that kept me going throughout this dissertation, especially after disappointing results or failed experiments.

And most of all, I am extremely appreciative to my friends and family. Zainab, your perspective is one I'll always need to hear. My family is a constant source of confidence and support—especially grandpa, who has never been short on life advice. Without the love and encouragement from my friends and family this dissertation would not have been possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

AMI	Advanced Metering Infrastructure
API	Application Programming Interface
AR	Augmented Reality
BUR	Building Use Ratio
eCPS	Energy Cyber Physical System
EUI	Energy Use Intensity
IDE	Integrated Development Environment
PV	Photovoltaic
RPS	Renewable Portfolio Standard
SDG	Sustainable Development Goal
SVM	Support Vector Machine
VIF	Variance Inflation Factor

SUMMARY

Of the 150 US cities that have adopted 100% clean energy resolutions, most are considerably far from achieving this goal. Concurrently, the rise of advanced and affordable sensors offering continuous monitoring of city infrastructure has directed research attention towards how data-driven approaches can help cities become 'smart' and achieve sustainability goals. As buildings account for the majority of energy consumption in cities, they have become a key focus for smart city initiatives. The influx of measurements on building energy and infrastructure at urban-scales add substantial complexity in handling of this information; an important area for research is developing approaches to translate this data into metrics that can be helpful for energy management and decision making at the community and city scale. Across the three studies in this dissertation, I take a multidisciplinary approach and draw on areas across data analytics, human-computer interaction, and public policy analysis to transform building energy data for improving community energy decisions. In the first study, I present a new approach for building energy benchmarking using building electricity smart meter data across the Georgia Tech campus. The results aim to support building portfolio owners and municipalities in identifying and prioritizing specific energy efficiency opportunities across a group of buildings. In a second study, I enhance the visibility and awareness of the same data through the development of a community-scale energy feedback system, and evaluate Georgia Tech community member understanding and reactions to having access to campus energy information. In a final study, I explore the impact of built infrastructure on renewable energy deployments across urban areas to inform urban planning design and

policy. The results of this work seek to contribute to research efforts within building energy efficiency fields and enhance our understanding of how advances in data science and computing can be connected to energy management and decision making practices. As cities strive to make substantial changes in their energy systems, emerging data sources may provide immense opportunities to make more effective and informed decisions; however, enabling this will require integration of new data science techniques with existing decision making practices. Continued connections between these two areas are likely to foster unique insights and pave the way for cities to attain a low-carbon future.

CHAPTER 1. INTRODUCTION

1.1 Global challenges, local action

Human extraction and consumption of natural resources across the past two centuries has enabled vast societal developments while causing a range of severe environmental and societal consequences. Air pollution threatens public health; more than 80% of people living in urban areas are exposed to poor air quality, which is linked to a multitude of respiratory diseases and health risks (World Health Organization 2016). Water contamination is degrading ecosystems; nearly 75% of the world's coral reefs are vulnerable, which support a substantial portion of aquatic life (Burke et al. 2011). Greenhouse gas concentrations are at an all-time high (World Meteorological Organization 2019), leading to warming that is associated with rising sea levels (Solomon et al. 2009), more extreme and frequent natural disasters (Mann and Emanuel 2006), and forced human migration (McAdam 2012), to name a few. In the midst of mounting evidence of the global impact of unrestrained human exploitation of resources, a large number of countries are recognizing that meeting the demands of future generations will require radical economic, societal, and technological transformations. For example, the majority of countries worldwide have made public commitments towards this shift—187 countries have ratified the Paris Agreement and 193 countries that have adopted the United Nation's Sustainable Development Goals (SDGs) agenda (United Nations 2019).

While national backing of the visions set forth by the Paris Agreement and UN SDGs is paramount, critiques have pointed out that voluntary, national-level commitments can lack accountability for achieving real results. Carbon reduction policies that contain mandates or enforcement mechanisms also often face political gridlock and can be difficult to implement. Tangible progress toward low-carbon societies will require efforts at more granular scales, such as the city or community-level (Van Der Schoor and Scholtens 2015). Cities and communities are uniquely positioned to facilitate sustainable transformations by enabling citizen participation and connecting global climate science issues closer to home. In effect, they are well positioned to implement context-specific initiatives that are more tailored and effective in reducing carbon emissions (Reckien et al. 2018).

1.2 Community energy initiatives and infrastructure data

One prominent area of focus for community sustainability initiatives is energy. Across the US, over 150 cities, ten counties, and seven states have adopted 100% clean energy goals, committing to powering their community entirely by clean energy by no later than 2050 (SierraClub 2019). Furthermore, a multitude of institutes and organizations are creating movements to congregate city leaders, share ideas, and encourage the creation of city-level climate action plans (C40 Cities 2020; We Are Still In 2020). The building sector is a priority within such movements. This is for good reason; buildings—including residences, offices, and restaurants, among others—are responsible for 40% of the energy use in the US, and make up 70% of electricity consumption (Energy Information Administration 2019). Approaches to reduce our reliance on carbon-intensive energy resources and achieve community energy commitments will require significant transformation in the building sector. New buildings will need to be built to high performance building standards, and existing buildings—with an average lifespan of 70 years—will need be renovated to operate at higher efficiencies while still meet building codes and occupant comfort needs (CBECS 2015). Municipalities have limited budgets and resources to improve the existing building stock; targeting specific behavioral, retrofit, and renewable energy opportunities that will have a high impact towards clean energy goals is essential. Recently, there has been momentum towards exploring the use of ‘big data’ sources to drive decision making and help target where to effectively allocate funds for building improvements (Zhou et al. 2016).

Recent pushes to collect data across buildings and the energy sector has been led by both public and private entities. Local governments have supported the public release of commercial building energy use and production data, namely through building benchmarking ordinances. As of June 2019, 32 cities and three states in the US have adopted building energy benchmarking disclosure policies, which mandate the tracking and reporting of annual building energy consumption and production information (Institute for Market Transformation 2019). The vision for this type of policy is threefold: (1) individual buildings will improve their energy management through intentional tracking with data, (2) cities will be able to more effectively allocate resources for efficiency improvement with city-scale building data, and (3) well-performing buildings will be rewarded by the market in terms of their value (Institute for Market Transformation 2020).

In addition to public data sources, private entities have pushed to increase monitoring capabilities in buildings. Advanced metering infrastructure (AMI) such as smart meters is proliferating in the building industry, being implemented by utilities, building managers, and consultants to monitor whole-building electricity use, as well as sub-metering of electricity loads, water flows, and other drivers of consumption. The big data analytics market within the energy sector is expected to have a compound annual growth rate of 10.22% between 2020 and 2025 (Mordor Intelligence 2019). Companies such as IBM, Cisco, and Siemens are leading efforts to connect this data to 'smart city' business models, offering services to deploy and maintain sensors, implement analytics and digitization to improve service delivery, and provide greater connectivity between government and its citizens (Cisco 2020; IBM 2020; Siemens 2020). This influx of data and digital management of infrastructure across the built environment at city-scales is unprecedented. Optimistically, such data opens the opportunity to improve our understanding of buildings and their operations, specific to a community in which the data is collected and processed. However, a multitude of concerns exist regarding the assumption that new data will generate new and better insights (Sutherland and Cook 2017).

1.3 Smart cities and data-driven decision making

Smart city initiatives often position data as a key resource, with the potential to enhance decision making. This emphasis on data-driven decision making is based on the premise that previously undiscoverable insights and new solutions can be garnered from sensor data that is becoming more affordable, granular, and ubiquitous across communities (Shelton et al. 2015). As our built infrastructure becomes increasingly complex, data is framed as a resource that can capture these complexities and enlighten objectively considered pathways to make better decisions (Hollands 2015; Zanella et al. 2014). A large body of research in computing and smart city domains has worked to establish theoretical frameworks for integrating disparate data sources, mining data for new insights, and creating analytical and predictive interfaces (e.g., Khan et al. 2015; Zhou et al. 2016). While more detailed information on our infrastructure has potential to inform decision making, visions presuming the inherent utility enabled by this data often lack purposeful integration into existing decision making practices connected to various stakeholders. Connecting data to people is essential, and remains a key challenge in smart city work today (Zuiderwijk et al. 2015).

There is a need in research to assess how emerging forms of building data are being used and identify potential ways to connect new forms of built infrastructure data to actual decision making processes for energy management and efficiency. From building managers to city planners to community residents, a wide array of actors have potential to better understand their community's energy use and make more informed decisions through access to information gleaned from new community-scale energy data. To enable this, there is opportunities for research to examine and develop approaches that transform emerging forms of community energy data, particularly open data, in ways that make it more understandable, engaging, and actionable to specific stakeholders.

1.4 Dissertation framing and structure

Community initiatives are an important part of global movements aimed at mitigating climate change. Energy efficiency and management efforts in the building sector have increasingly concentrated on the production and collection of data in order to inform decisions, however a disconnect exists between the multitude of data collected and how stakeholders currently navigate energy management decisions. New advanced metering infrastructure data has potential to be particularly useful as it can offer granular insights in a localized and community context; however effective utilization of this data is unlikely to be achieved without integration into existing decision making practices and engagement with specific stakeholders. This dissertation aims to contribute to the fields of building energy analytics and visualization by increasing the usability of emerging forms of building energy data. It entails three studies where I harness emerging sources of building energy data and develop approaches to inform building energy decision making at community and city scales.

In Chapter 2, I present a new approach for building energy benchmarking using smart meter data and Georgia Tech's campus as a research testbed site. While traditional annual building energy benchmarks offer useful techniques for comparing energy performance across a group of buildings, they are limited in their ability to give insight about specific retrofit opportunities. To fill this gap, I developed an approach that leverages smart meter data, which is becoming increasingly available to cities and building portfolio owners. This benchmarking approach uses top-down statistical techniques and generates more specific insights facilitated through granular smart meter data. The results can support building portfolio owners and municipalities in identifying specific energy efficiency opportunities across a community-scale of buildings. The article in Chapter 2 is published in ASCE's *Journal of Management in Engineering* (Francisco et al. 2020).

In Chapter 3, I describe the development of a community-scale energy feedback system built for community members of the Georgia Tech campus. The notion that citizens should take on more active engagement in their day to day lives to reduce carbon emissions has long been regarded as essential. Smart city open data initiatives also claim that new, open sources of data bring new opportunities for citizen involvement and engagement with governments. However, there is a need to transform this data for the intended users. In this study, I leverage the same data collected in Chapter 2 and apply novel visualization techniques, such as augmented reality, to spatially connect users with building performance and enhance community member engagement with campus energy systems. Real user understandings and perceptions of the campus-scale energy data are analyzed to examine the potential of community energy feedback in informing community member decision making and energy actions. The article presented in Chapter 3 is published as a journal article in the *Journal of Applied Energy* (Francisco and Taylor 2019a).

In Chapter 4, I empirically assess the relationship between urban form and solar PV adoption decision making. Recognizing the long-lasting impact of infrastructure decisions and how cities are currently transforming and infilling their infrastructure substantially, I investigate how this change may impact the viability of clean energy technologies in cities. While in Chapter 2 I explore relationships between energy consumption activities across different building characteristics, this study examines relationships between energy production activities across different building characteristics. Specifically, this study focuses on rooftop solar suitability. The results can inform city planners and urban designers in how important rooftop solar suitability's and, more broadly, urban form's role is in solar PV adoption decisions. The article presented in Chapter 4 has been developed into a journal paper and will be submitted to a journal that publishes contributions at the intersection between energy systems and urban design.

Finally, Chapter 5 includes a discussion of the overarching contributions of this work in the building energy efficiency, analytics, and visualization fields. This work has drawn from a multitude of areas including energy analytics, human-computer interaction, and public policy. This chapter draws connections between these areas and recommends areas for future work based on the findings of this dissertation.

CHAPTER 2. SMART CITY DIGITAL TWIN-ENABLED ENERGY MANAGEMENT: TOWARD REAL-TIME URBAN BUIDLING ENERGY BENCHMARKING[1]

2.1 Introduction

In cities, prominent challenges such as urbanization and rising greenhouse gas emissions have sparked efforts to make cities 'smarter' (Pierce and Andersson 2017). As buildings account for the majority of energy consumption in cities, and because of their high potential for energy conservation through retrofits or operational improvements (IPCC 2007), they have become a key focus for smart city initiatives (Baxter et al. 2011). At the intersection of smart cities and building energy efficiency lies the opportunity for real-time intelligent planning and urban energy management (Hastak and Koo 2016). Smart city digital twins, a recent endeavor to create a digital replica of city infrastructure linked to real-time city data, is envisioned to improve city monitoring, control, and decision making through enhanced visualization and interaction with city data (Mohammadi and Taylor 2017). Smart city digital twins are intended to capture and incorporate urban complexities across time and space through streamed data; given the increasing availability of building performance data at urban scales (BuildSmart DC 2017), smart city digital

[1] This chapter was published as a journal article in the ASCE Journal of Management in Engineering with Neda Mohammadi and John E. Taylor as the co-authors. The citation for the journal article is as follows: Francisco, A., Mohammadi, N., and Taylor, J.E. (2020). "Smart City Digital Twin-Enabled Energy Management: Toward Real-Time Urban Building Energy Benchmarking." *Journal of Management in Engineering*, 36(2).

twins are a promising platform for building portfolio performance assessment and urban energy management (i.e., digital twin-enabled energy management).

Concurrently, policies aimed at transitioning cities to more sustainable, energy-efficient urban areas are only growing. As of March 2019, over 90 cities, ten counties, and two states in the U.S. have committed to consuming energy entirely from renewable energy sources by no later than 2050 (SierraClub 2019). Other policies such as building benchmarking ordinance requirements are requiring public release of whole building energy consumption and production data for individual buildings at community and city scales (BuildingRating 2019; BuildSmart DC 2017). Harnessing the potential of such data, made available through large investments in smart infrastructure, is critical to fulfill greenhouse gas emission reduction commitments (Zuo et al. 2013) and to strive towards digital twin-enabled smart city energy management.

Traditional ways of accomplishing building portfolio assessments across large-scales include building energy benchmarking, which is typically conducted on an annual basis (Borgstein et al. 2016). However, such metrics do little to inform specific efficiency opportunities to target or support real-time management of energy efficiency. The availability of smart meter data across a community of buildings can enable the construction of benchmarks developed at finer temporal scales and across specific time segments, or what we define as temporally segmented building energy benchmarks. As many buildings require different levels of energy consumption based on the time of day or week, temporally segmented building energy benchmarks have potential to provide a more accurate measure for building efficiency across a group of buildings (Francisco et al. 2018a; Roth and Jain 2018). In this paper, we leverage smart meter data for a community of buildings to benchmark energy consumption at different temporal periods and determine that temporally segmented building energy benchmarks integrated with digital twins can help detect previously undiscoverable insights and identify more specific time-driven strategies for near real-time urban energy management.

2.2 Background

2.2.1 Energy performance assessments in existing buildings

Energy performance assessments are commonly used to assess the energy performance of existing buildings, and can be divided into two categories: building energy classification and energy diagnosis (Wang et al. 2012). Building energy classification is applied across a group of buildings and adopts methodologies that standardize each building's total energy consumption based on its characteristics, enabling comparison of building energy efficiency levels between buildings with different characteristics (e.g., floor area, space use) (Pérez-Lombard et al. 2009). One of the most commonly used standardization processes within energy classification is *building energy benchmarking*, and is a top-down approach applying statistics or machine learning algorithms (Buck and Young 2007; Kavousian and Rajagopal 2014; Zhang et al. 2015) that classify whole-building energy efficiency levels on an annual basis (Borgstein et al. 2016; Kinney and Piette 2002). Such macro-level metrics are generally easily understood and aim to connect building owners, policymakers, and the public with the information necessary to identify poor performing buildings and motivate stakeholders to implement energy efficiency improvements (Wang et al. 2012). In general, these approaches require little data on building technology (Zhao et al. 2017), or the physical characteristics of a building (Hong et al. 2012; Li et al. 2014), which is advantageous in large scale studies across a community or city where the availability of such data in most existing buildings is limited. While benchmarking techniques help identify overall building performance, their findings are not specific enough to guide operational or physical improvement recommendations for a building (Borgstein et al. 2016). To identify the root-causes of building inefficiencies, energy diagnosis methodologies are often required (Borgstein et al. 2016).

In contrast to building energy benchmarking, energy diagnosis methodologies use bottom-up approaches to help identify where energy inefficiencies exist in a building. Common methodologies for diagnosis include energy simulations and engineering calculations (Borgstein et al. 2016; Burman et al. 2014). Most commercially available energy modeling software requires building geometries, physical properties, thermal zones, and operational inputs to simulate building energy consumption (Burman et al. 2014). Additionally, engineering calculations methodologies, often accomplished through energy audits, use building system specifications and operational schedules to predict energy consumption and efficiency levels of separate building systems (Burman et al. 2014). Diagnosis methodologies provide results that are easily translated into specific steps to improve building energy performance, and thus overcome shortcomings of the energy benchmarking methodologies mentioned above. However, they are limited for the following reasons. First, while research has made substantial progress in the reconciliation between estimated energy consumption from bottom-up methodologies with measured consumption, their accuracy are not yet consistently reliable, and the reconciliation process is time intensive, computationally expensive, and requires building science expertise (Wang et al. 2012). Second, bottom-up approaches rely on extensive and accurate data collection of building systems on site or through detailed review of building plans (Borgstein et al. 2016). As municipalities are often constrained by budgets and manage diverse buildings, the extensive time and experience required to comprehensively assess city-wide building performance using bottom-up methodologies is not practical.

Given the advantages and disadvantages of energy benchmarking and diagnosis approaches, community or city level building energy assessments could be improved by providing more actionable and system-specific efficiency indicators (i.e., the results of energy diagnosis) while still applying methodologies that do not require extensive data collection, time, or expertise (i.e., the methodologies of energy benchmarking). One promising research avenue aiming to develop more actionable results while leveraging the scalability of benchmarking methods is applying smart meter data to examine building energy benchmarks segmented by time. As real-time energy management tools become integral to energy efficiency decision making (Kitchin et al. 2015; Ramachandra et al. 2018), energy benchmarks segmented by time have immense potential to further the efficacy of digital twin-enabled energy management platforms if they can provide more insights, value, and frequent feedback compared to their annual counterparts.

2.2.2 The temporal dimension of building energy performance

While building characteristic data (e.g., wall insulation levels, HVAC equipment types, number of appliances) is challenging to collect at scale, energy data recording electricity use at granular levels (meter readings less than once per hour) is becoming increasingly accessible through advanced smart metering infrastructure (EIA 2018). The availability of smart meter data has spurred a variety of research assessing how this data, combined with statistical or machine learning algorithms, can support a wide range of applications including energy load analysis, forecasting, and management (Wang et al. 2018). Broadly defined as smart meter data analytics, such research enables near real-time analyses of energy use (Wang et al. 2018), which was previously not feasible using traditional energy meters. Previous literature has used smart meter data extensively across various applications to further real-time analytics. For example, this data has been used to detect anomalous consumers in real-time using decision trees and support vector machine (SVM) classifiers (Jindal et al. 2016), understand energy behaviors of commercial building occupants using k-Nearest Neighbors classifier (Rafsanjani et al. 2017), and optimize real-time energy pricing using innovative clustering techniques (Joseph and Erakkath Abdu 2018). In general, smart meter data analytics applications have focused on consumer segmentation, prediction, and demand response applications. Few studies have examined the potential of smart meter data analytics applied to building benchmarking and performance applications, particularly for real-time use across larger scales of buildings.

A few recent studies have integrated smart meter data into building benchmark analyses (ElYamany et al. 2017; Francisco et al. 2018a; Roth and Jain 2018). The general approach involves leveraging the granularity offered by smart meter data to segment building efficiency levels by different time periods and help enlighten more specific areas for efficiency improvement, which can be applied across a large scale of buildings. This approach stands in contrast to traditional building energy benchmarks, which are calculated on an annual basis and offer little insight into how to improve building performance. Grolinger et al. (2017) calculated building energy benchmarks during time periods when events occurred in two sports arenas to identify the most and least efficient events. The purpose of these benchmarks was to help identify efficiency opportunities and where to prioritize efficiency efforts across the two sports arenas. In a comparable analysis, Roth and Jain (2018) expanded the scope of buildings included to assess 500 K-12 school buildings. They benchmarked energy use segmented by operational and non-operational time periods and suggested potential targeted areas for efficiency opportunities based on visual analysis of these metrics. The results of both of these studies suggest differences exist between temporally segmented building energy benchmarks, and the likely potential for these to enlighten efficiency opportunities across a group of buildings. However, these analyses are based on the assumption that temporally segmented benchmarks differ systematically from their annual counterparts, while the magnitude, distribution, and statistical significance of the differences between temporally segmented and annual benchmarks has yet to be examined, which may compromise the validity of such an approach.

With research only beginning to examine the development and utility of temporally segmented energy benchmarking, this paper seeks to contribute to this work by investigating the research question: do temporally segmented building energy benchmarks differ from their non-temporally segmented counterparts? If so, how can temporally segmented building energy benchmarks be used to improve energy management across a portfolio of buildings? To evaluate this, we use smart meter electricity data across a community of buildings to develop daily benchmarks that are segmented by strategic time periods and statistically evaluate their deviations from a control group, their non-temporally segmented counterparts. Following this analysis, we discuss in detail the practical implications of these findings, particularly through a smart city digital twin lens.

We elaborate on how these results establish new capacities for digital twin-enabled energy management through several examples that show the potential of temporally segmented energy benchmarks to better inform energy efficiency decision making by enabling: (a) identification and prioritization of specific retrofit strategies, and (b) near real-time building energy management. Both of these implications propel the development of digital twin-enabled energy management systems, which require informative metrics to be successful in supporting building operators or portfolio owners with energy decisions.

2.3 Methods

The Georgia Institute of Technology (GT) university campus was selected as a testbed to quantify building energy efficiency scores' temporal variation within a community. University campuses have diverse and dynamic operations, consisting of a combination of offices, laboratories, recreation, health, food, retail, and classroom facilities that are comparable to the operations of a small town or community (Klein-Banai and Theis 2011). The data scope of this analysis covers building-level electricity consumption for 38 buildings on the GT campus. While the buildings in this sample have diverse functions, they all have heating and cooling provided by a district water loop. Buildings with individual electric cooling or heating systems were eliminated from the sample to avoid bias in efficiency scores due to unequal end uses. Average power is recorded in 15-minute increments and the experimental dataset ranges from September 26, 2015 to September 25, 2016. In the following sub-sections, the methods used to segment building energy consumption by time period, compute daily energy benchmarks for each building, and conduct a series of hypotheses tests are described.

2.3.1 Data segmentation by temporal period

To establish building energy benchmarks for each building across each temporal period, electricity use for each building was first segmented by the following time periods: occupied periods during the school year (A), unoccupied periods during the school year (B), occupied periods during the summer (C), unoccupied periods during the summer (D), peak summer periods (E), and the total period (Total). While additional temporal periods could be selected and assessed, the periods listed above were selected based on their alignment with operational shifts that buildings commonly undergo and are supported by the literature in having a high potential to enlighten efficiency opportunities. For example, studies have found that substantial energy waste occurs in buildings during unoccupied periods due to misaligned operational schedules (Gul and Patidar 2015; Masoso and Grobler 2010). Knowing this, it is likely helpful to differentiate between efficiency levels during occupied and unoccupied in targeting efficiency opportunities. Building operations also can shift seasonally. One study found that annual energy efficiency scores for university buildings were skewed due to significant operational shifts during the summer months, and recommended separating summer months from annual efficiency scores to uncover actual efficiency levels (Tu 2015). Last, energy consumption during summer peak demand periods is a pressing concern for facility managers due to utility peak demand charges (Neufeld 1987). Specific retrofit opportunities exist to reduce energy demand during summer peak periods, such as improving air conditioner efficiency (Yarbrough et al. 2015).

Table 1 outlines the times and days included in each temporal period specified above, which encompass a one year period in total. The start of the one year period was selected strategically to minimize the amount of data gaps in the dataset. In determining occupancy states, measured occupancy levels were not available for every building in the sample and building occupancy states were assumed to be occupied between 8AM and 8PM each weekday. These hours reflect when building entrances on campus are typically open/unlocked (~8AM), versus when they start to require key access (~8PM). Hours after 8PM and before 8AM were assumed to have low to no occupancy and will be referred to as the unoccupied state in this paper. Weekends were excluded from the analysis because occupancy states during weekends are less consistent, and therefore estimates would be less reliable. Additionally, the two week period encompassing the winter break (between December 19, 2015 and January 3, 2016) was removed from the analysis due to unknown occupancy states. The seasonal shifts were divided between the school year and summer, as defined by the Georgia Tech school calendar for the 2015 – 2016 school year. The summer peak demand time range was based on when Georgia Power, the power supplier of the campus, charged customers peak billing demand rates during the summer.

Table 1 – Temporal Period Details

Temporal Period	Days/times	Total Number of Days
Occupied during school year (A)	*Days:* 9/26/15 – 5/7/16, 8/21/16 – 9/25/16 *Times:* 8AM – 8PM (M-F)	174
Unoccupied during the school year (B)	*Days:* 9/26/15 – 5/7/16, 8/21/16 – 9/25/16 *Times:* 8PM – 8AM (M-F)	174
Occupied during the summer (C)	*Days:* 5/8/16 – 8/20/16 *Times:* 8AM – 8PM (M-F)	75
Unoccupied during the summer (D)	*Days:* 5/8/16 – 8/20/16 *Times:* 8PM – 8AM (M-F)	75
Peak summer (E)	*Days:* 9/26/15 – 9/30/15, 6/01/16 – 9/25/16 *Times:* 2PM – 7PM (M-F)	86
Total	*Days:* 9/26/15 – 9/25/16 *Times:* 12AM – 11:59PM	365

Next, the energy data for each time period were aggregated from 15-minute interval data to daily energy use values, referred to in this paper as the *temporally segmented* data, in order to perform energy benchmarking on a daily basis. To aggregate energy data to the daily level, for each day the average of the 15-minute interval energy data was calculated. This procedure was followed for temporal periods A, B, C, D, and Total. Because peak period charges are based on the maximum kW in a 30 minute period, the data during the peak summer period (E) was averaged using a 30-minute running average, and then the maximum value for the day was selected. The total period energy data represents the average electricity use across all 24 hours of each day and serves as the control. After following energy benchmarking procedures, detailed below, energy efficiency scores for each building across each time period were generated. Temporally segmented and total period building energy daily efficiency scores were compared for each building, meaning

38 comparisons are conducted for each temporal period. The hypotheses tested are as follows:

- Hypothesis A (1-38): A building's daily efficiency scores during the total period compared to its daily efficiency scores during *occupied periods in the school year* are not the same.
- Hypothesis B (1-38): A building's daily efficiency scores during the total period compared to its efficiency scores during *unoccupied periods in the school year* are not the same.
- Hypothesis C (1-38): A building's daily efficiency scores during the total period compared to its efficiency scores during *occupied periods in the summer* are not the same.
- Hypothesis D (1-38): A building's daily efficiency scores during the total period compared to its efficiency scores during *unoccupied periods in the summer* are not the same.
- Hypothesis E (1-38): A building's daily efficiency scores during the total period compared to its efficiency scores during *summer peak demand periods* are not the same.

2.3.2 Efficiency score development

The purpose of building benchmarking is to generate building efficiency scores that enable more accurate comparisons of energy efficiency between buildings. Early benchmarking methods created simple ratio metrics, such as energy use per area or energy use per occupant, which normalized energy use based on a single building characteristic

(Pérez-Lombard et al. 2009). Later, more complex approaches such as statistical or machine learning methods were introduced, which attempted to normalize energy use based on more than one building characteristic. Using such approaches, an efficiency score is generated by adjusting a building's energy use to account for multiple building characteristics simultaneously. In this study, we adopted the regression-based methodology developed in Chung et al. (Chung et al. 2006). Such an approach is similar to methodologies commonly used by industry benchmarking applications such as the ENERGY STAR® score (Borgstein et al. 2016; Shrestha and Kulkarni 2013). The following paragraphs detail the steps taken to apply this methodology to generate building efficiency scores for each day within each temporal period.

First, the dependent and independent variables for the benchmark model were defined. The independent variables represent explanatory variables of energy consumption, which were used to normalize energy use across different buildings (Chung et al. 2006). These variables encompass building characteristics that are outside the control of the building operators or occupants, to normalize building energy use based on the building features that are unlikely or infeasible to change. For example, different building space types (e.g., lab and office spaces) have different energy requirements for operation (Park et al. 2016). Normalizing building energy use by the area of different space types can enable more accurate comparison of the energy use of buildings with different functions. Other inflexible building characteristics (i.e., characteristics that cannot be easily changed by management or occupants), were collected from a publicly accessible database from the university's Capital Planning and Space Management group. All explanatory variables and the dependent variable used in the model are outlined in Table 2. Space type areas were converted to Building Use Ratios (BURs), ranging between 0 and 1, by dividing the area of the space type by the total building floor area (Park et al. 2016). In addition, the dependent variable for each temporal period was divided by building floor area to generate Energy Use Intensities (EUI), which is the primary unit for energy benchmarking analyses as building floor area is highly correlated with energy use (Sharp 1995). Similar to the approaches of previous benchmarking studies (Buck and Young 2007; Chung 2012; Park et al. 2016), floor area was selected to also be included as an independent variable. Figures with the distributions of the explanatory variables can be referred to in the Appendix (Figure 16).

Table 2 – Explanatory and Dependent Variable Descriptions

Variable	Feature	Unit
Independent Variables	Floor area	ft^2
	Building age	Years
	Years since renovation	Years
	Number of floors	Floors
	Percent renovated	%
	BUR: Laboratory	%
	BUR: Office	%
	BUR: Mechanical	%
	BUR: General	%
	BUR: Circulation	%
	BUR: Service	%
	BUR: Supply	%
	BUR: Classroom	%
	BUR: Study	%
	BUR: Special	%
Dependent Variable	Daily Average EUI (Total period, A-E temporal periods)	$kWh/ft^2/day$

Next, temporally segmented daily EUIs for each building were normalized by the identified explanatory factors using a multivariate linear-regression approach, similar to Chung et al. (Chung et al. 2006). The first step of the normalization process is to create a model to quantify the relationship between the building explanatory factors and the EUIs. A regression model was created for each day within each temporal period. For example, 174 models were created for temporal period A (see Table 1). Several transformations were made to the independent and dependent factors. To account for skewed distribution characteristics, EUI and floor area model inputs were log-transformed. The explanatory variables in Table 2 were rescaled to have a mean of zero and standard deviation of one in order to aid the interpretation of the regression coefficient results. The rescaled explanatory variables served as regression model inputs $(x_1, \dots , x_p)$. For each day within each temporal period, the regression model took the following form:

$$EUI = a + b_1 x_1^* + \dots + b_k x_k^* + \varepsilon \qquad (1)$$

where a is the intercept, ε is the error term, $\mathrm{x}_1^*, \dots , \mathrm{x}_\mathrm{k}^*$ are the significant explanatory variables (where k ≤ p), and $b_1, \dots , b_k$ are the model coefficients. Forward selection was applied to identify the significant explanatory values included in the model, with the Akaike Information Criterion providing the basis for variable selection. Based on the regression model results for a particular day, the daily EUI for a building can be normalized by the following:

$$EEUI_{norm} = EUI_o - EUI + a \qquad (2)$$

where EUI_o is the measured EUI of the building for that day, EUI is the predicted EUI based on the model coefficients from Equation 1, a is the model intercept in Equation 1, and EUI_{norm} is the building's normalized EUI. Equation 2 is equivalent to calculating the residual for a building ($\mathrm{EUI_o} - \mathrm{EUI}$) and adding this to the model intercept. The model intercept a represents the EUI for the average building in the dataset, given that the explanatory variables are rescaled with a mean of zero. To illustrate the normalized EUI calculation, if the expected EUI for a building (EUI) is less than the actual EUI (EUI_o), this difference will be added to the average EUI across the buildings (a), leading to a higher normalized EUI (EUI_{norm}). A regression model was created for each day within each temporal period, thus the explanatory variables, coefficients, and intercept values in Equations 1 and 2 changed from day to day across temporal periods.

After normalizing each building's EUI, the EUI_{norm} values were rescaled between 0 and 1 for each day. These represented the efficiency scores, where 0 is the least efficient and 1 is the most efficient. The result was a distribution of daily efficiency scores for each building within each temporal period.

To compare the efficiency score distributions and test the hypotheses, the Wilcoxon signed-rank test is used, which is a non-parametric test that accepts ordinal data, such as energy efficiency scores. The Wilcoxon signed-rank test is a paired test, thus, a building's efficiency score during the total period was compared to the temporally segmented period for the same day. As the total period had more days compared to the temporally segmented periods, efficiency scores were only included in the statistical analysis if they were both computed on the same day. Each hypothesis was tested for each building, introducing the multiple comparison problem, which increases the probability of committing a Type I error

(Bretz et al. 2011). Therefore, adjusted *p*-values were computed to control for family-wise errors using the Holm procedure. The Holm procedure is a more powerful modification of the Bonferroni correction and was used due to the Bonferroni often being regarded as too conservative when a large number of tests are conducted (Holm 1979). Confidence intervals of 95% or greater indicated statistically distinct efficiency score medians. Efficiency score development and statistical analysis were calculated using R (R Core Team 2017).

2.4 Results

The regression model coefficients for each day were used to normalize building daily EUI and generate energy efficiency scores. To test Hypothesis A1, the distribution of daily efficiency scores for Building 1 during *occupied hours in the school year* was compared to the distribution of daily efficiency scores for the same building during the *total period* using the Wilcoxon signed-rank test (i.e., a paired difference test). This is repeated for all 38 buildings in Hypothesis A. Next, this process is repeated for the remaining temporally segmented periods. Adjusted p-values, to account for the multiple comparison problem, are displayed in Table 3. At the 95% confidence interval, there was enough evidence to reject the null hypothesis for 34 buildings in Hypotheses A, 32 building in Hypotheses B, 31 buildings in Hypotheses C, 30 buildings in Hypotheses D, and 32 buildings in Hypotheses E. The final column lists the total number of times the null hypotheses were rejected for each building. Almost 75% of the buildings (n = 28) had a statistically distinct distribution of efficiency scores in four or more of the examined temporal periods compared to the total period. All of the buildings had a statistically significant difference for at least one temporal period. The p-values equivalent to 1 is the result of applying the Holm procedure, which adjusted the p-values to be more conservative (in terms of Type I errors) in order to account for the multiple hypothesis problem.

Table 3 – Adjusted p-values for Hypotheses A, B, C, D, and E

Building	A	B	C	D	E	Total Sig. Cases
1	<0.001***	<0.001***	<0.001***	<0.001***	<0.001***	5
2	<0.001***	<0.001***	<0.001***	<0.001***	<0.001***	5
3	<0.001***	<0.001***	<0.001***	<0.001***	0.002**	5
4	<0.001***	<0.001***	<0.001***	<0.001***	<0.001***	5
5	<0.001***	<0.001***	<0.001***	<0.001***	<0.001***	5
6	<0.001***	0.005**	0.048*	<0.001***	<0.001***	5
7	<0.001***	<0.001***	<0.001***	<0.001***	<0.001***	5
8	<0.001***	<0.001***	<0.001***	<0.001***	<0.001***	5
9	0.021*	<0.001***	<0.001***	<0.001***	0.03*	5
10	<0.001***	<0.001***	<0.001***	<0.001***	<0.001***	5
11	<0.001***	<0.001***	<0.001***	<0.001***	<0.001***	5
12	<0.001***	<0.001***	<0.001***	<0.001***	<0.001***	5
13	<0.001***	0.034*	<0.001***	<0.001***	<0.001***	5
14	<0.001***	<0.001***	<0.001***	<0.001***	<0.001***	5
15	<0.001***	<0.001***	<0.001***	0.004**	<0.001***	5
16	<0.001***	<0.001***	0.006**	<0.001***	0.026*	5
17	<0.001***	<0.001***	<0.001***	<0.001***	<0.001***	5
18	<0.001***	<0.001***	<0.001***	<0.001***	<0.001***	5
19	<0.001***	<0.001***	<0.001***	<0.001***	<0.001***	5
20	<0.001***	<0.001***	<0.001***	<0.001***	<0.001***	5
21	<0.001***	<0.001***	<0.001***	<0.001***	<0.001***	5
22	0.005**	0.005**	<0.001***	<0.001***	0.728	4
23	1	<0.001***	<0.001***	0.001**	<0.001***	4
24	<0.001***	0.005**	<0.001***	1	<0.001***	4
25	<0.001***	<0.001***	<0.001***	0.085	<0.001***	4
26	<0.001***	1	<0.001***	<0.001***	<0.001***	4
27	1	<0.001***	<0.001***	<0.001***	0.002**	4
28	<0.001***	<0.001***	0.959	<0.001***	<0.001***	4
29	0.001**	0.309	0.959	0.001**	<0.001***	3
30	<0.001***	0.473	0.003**	<0.001***	0.728	3
31	<0.001***	0.003**	0.009**	1	0.728	3
32	<0.001***	<0.001***	0.312	1	0.011*	3
33	0.006**	<0.001***	0.959	<0.001***	1	3
34	<0.001***	0.004**	0.006**	1	0.292	3
35	<0.001***	0.075	<0.001***	1	<0.001***	3
36	0.687	<0.001***	0.445	<0.001***	1	2
37	<0.001***	0.063	0.443	1	<0.001***	2
38	0.773	1	0.765	0.264	0.03*	1
Total Sig. Cases	**34**	**32**	**31**	**30**	**32**	

*Statistical significance: ***($p < 0.001$), **($p < 0.01$), *($p < 0.05$)*

A summary of the magnitude of the differences between total and temporally segmented efficiency scores for all statistically significant buildings is shown in Figure 1. Among the significant buildings, the mean absolute difference between the total and temporally segmented efficiency scores during temporal period A was 0.079, while the maximum absolute difference was 0.483. This implies that across the buildings with statistically significant efficiency scores differences, daily efficiency scores during occupied periods in the school year differed by an average of 7.9% from the total efficiency score, and by as high as 48.3%. These magnitudes changed by temporal period, summarized in Figure 1. Notably, these differences are in comparison to the total period, which is representative of the average efficiency score between all temporal periods. Thus, computing the differences between efficiency scores of two temporally segmented periods would many times result in even larger differences.

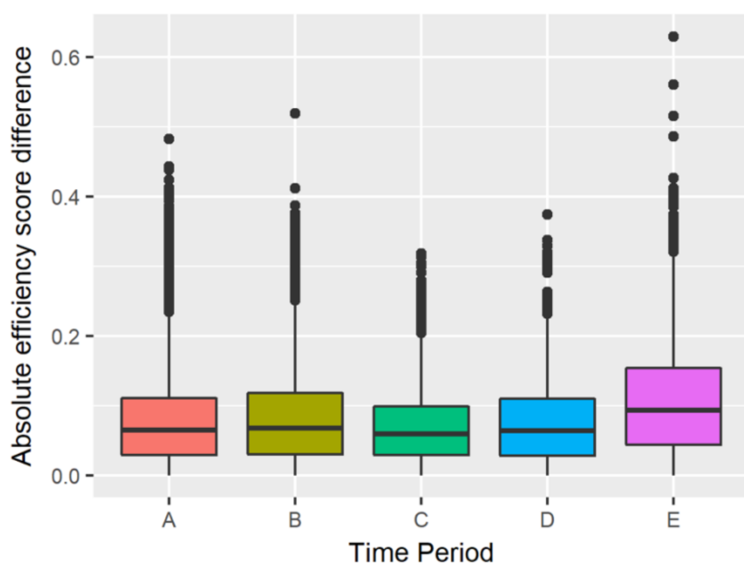


Figure 1 – Absolute difference between total period efficiency scores and temporal period efficiency scores for statistically significant buildings

2.5 Discussion

This study developed building energy benchmarking scores segmented by strategic time periods and statistically assessed how they vary from conventional, total benchmarking scores. Daily scores that vary systematically from total scores can enable more real time and more informative predictions to guide operational decision-making. Leveraging individual building smart meter data across a portfolio of buildings enabled the development of daily, temporally segmented benchmarking scores, and the results of the statistical analysis showed that temporally segmented building energy benchmarks were significantly distinct from their total counterparts for the vast majority of buildings in the sample (between 30 and 34 out of the 38 buildings in the portfolio). Previous studies have pointed out how conventional building energy benchmarks are limited in their ability to help target areas for efficiency improvement (Borgstein et al. 2016). While recent work has leveraged smart meter data to explore the potential of temporally segmented energy benchmarks in gaining more specific efficiency insights (ElYamany et al. 2017; Francisco et al. 2018a; Roth and Jain 2018), the deviations between this novel technique and their conventional counterparts had not previously been statistically assessed. This study contributes to emerging work examining the temporal dimensions of energy benchmarking (ElYamany et al. 2017; Francisco et al. 2018a; Roth and Jain 2018) by assessing the statistical significance and magnitude of differences between temporally segmented and conventional benchmarks for a community of buildings. In addition, this work furthers smart meters analytics research by documenting how this area can be integrated with and applied to building energy benchmarking methods. This is a critical step to understand deviations in building performance throughout the day and year relative to a community of buildings, which importantly has several practical implications for urban energy management.

The results of daily and temporally segmented benchmarks detect performance variations across time and have the potential to support with targeting, prioritizing, and managing individual building efficiency opportunities across a large geographic scale of buildings. While digital twin-enabled energy management platforms are envisioned to stream energy data sources (e.g., smart meter data) and are conceptually a promising platform to support cities with building portfolio performance assessments and urban energy management (Mohammadi and Taylor 2017), the construction of dynamic and real-time metrics that transform smart meter data into useful information is integral element for digital twin-energy management to be successful. A smart city digital twin energy management platform built around temporally segmented building energy benchmarks (example in Figure 2) offers the potential for managers to identify and prioritize specific retrofit strategies and detect near-real time deviations in building efficiency in the context of the performance of the entire building portfolio. The following two sections provide specific examples of how temporally segmented building energy benchmarks facilitate both energy efficiency prioritization and near real-time decision making.

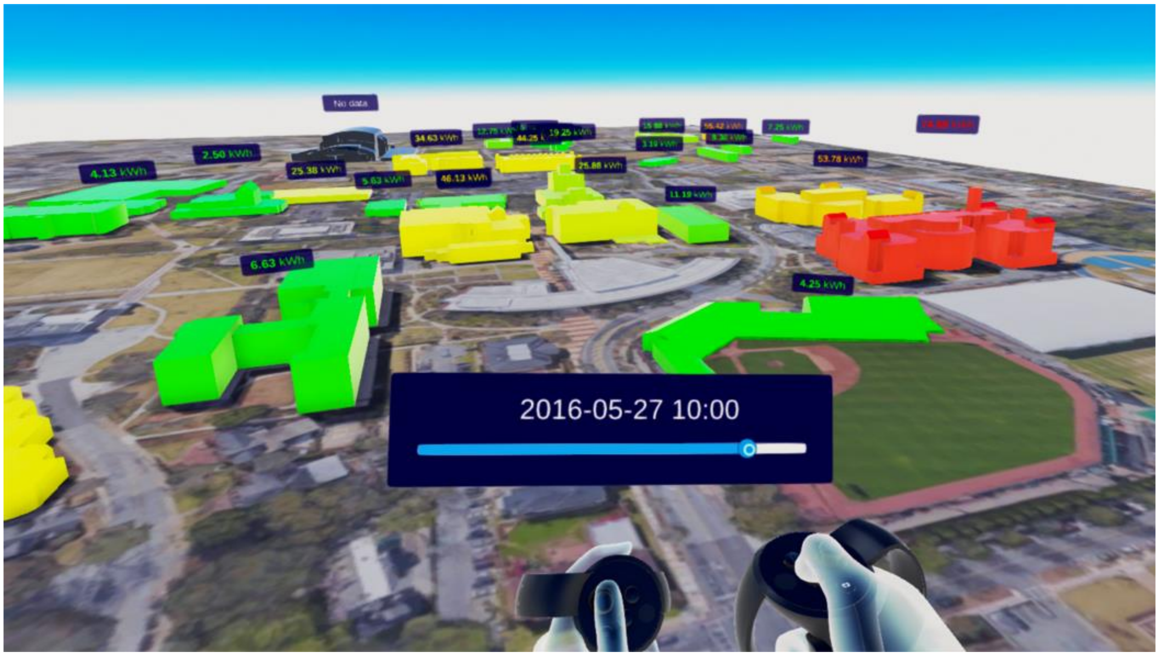


Figure 2 – Digital Twin-enabled Energy Management Platform

2.5.1 Prioritization of specific retrofit strategies across buildings

Temporal fluctuations in efficiency scores can enlighten how to prioritize building efficiency improvements. Figure 3 illustrates the difference between total and summer peak daily efficiency scores across the temporal state ‘summer peak demand’ for four buildings. Building 18 is consistently green across the summer, meaning the total daily efficiency scores were consistently higher (i.e., more efficient), than the peak summer efficiency scores. This distinction is necessary when prioritizing efficiency improvements, as specific measures are appropriate for reducing energy use during summer peak periods, such as increasing air conditioner efficiency or peak-load shifting (Koomey and Brown 2002). If a building manager were only considering the total efficiency score for Building 1, this building would appear to be more efficient than its actual performance during summer peak demand hours. More specifically, the total efficiency scores mask the peak efficiency score with differences up to 34.1%. This building's efficiency score rank during summer peak demand periods can help building managers decide whether to invest in peak demand reduction improvements with this building.

On the other hand, different trends in summer peak demand efficiency scores were seen in Buildings 20, 6, and 8. Building 20 is predominantly purple, demonstrating that the summer peak demand efficiency scores were often more efficient than the total efficiency score. This indicates that this building is more efficient during summer peak demand periods, compared to the total period, and likely should have lower prioritization for resources allocated to reduce peak demand. Alternatively, Building 8 transitions from predominantly green days to predominantly purple days, while Building 6 has more mixed variations. These trends show that energy performance during peak periods relative to total periods is changing sharply from day to day or week to week. In buildings where these changes are more extreme (e.g., indicated by the darker shade of the color), operational measures such as reviewing the building management system or other automated controls should be investigated to determine if the programmed operations still reflect the actual building conditions.

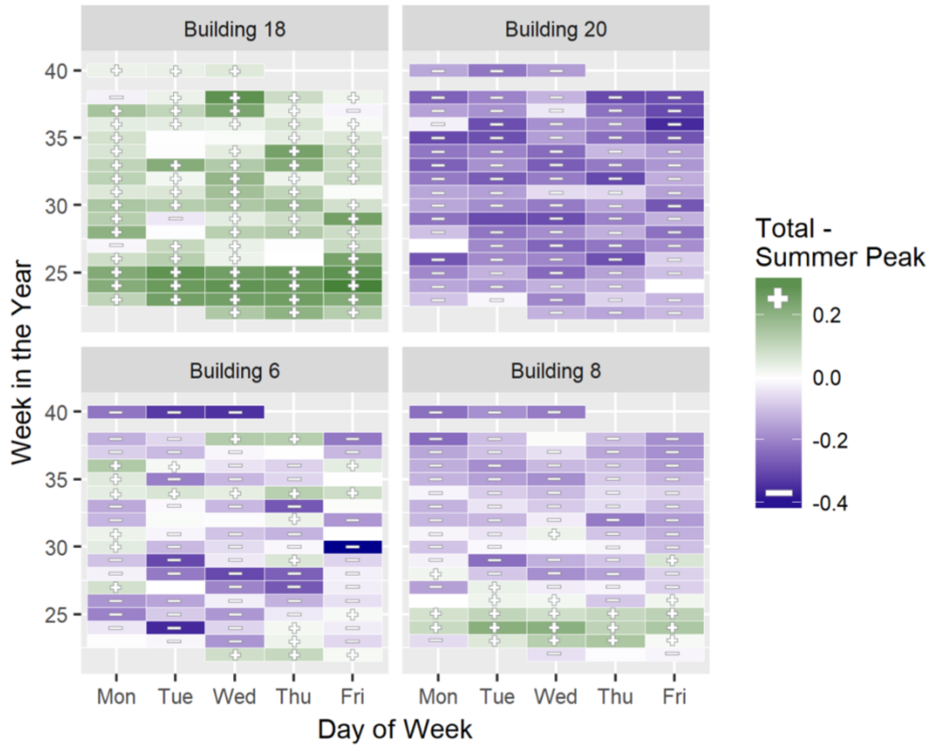


Figure 3 – Difference in Daily Efficiency Scores for Summer Peak Periods

2.5.2 Near real-time energy management across buildings

The previous example considered deviations between the total and temporally segmented efficiency scores to compare efficiency opportunities between buildings. Figure 4 presents the raw efficiency scores across temporal periods to show how building performance is changing over time within buildings. This can help generate insights with several uses, including identifying buildings with sudden changes in performance and buildings with consistently low levels of performance, and helps demonstrate how temporally segmented benchmarks can support more real-time energy management. Buildings with sudden changes in performance may simply require a review and update to operational controls, while buildings with continuously low performance may require investment in more capital-intensive upgrades. Of note, lines for only three segmented temporal periods are shown in Figure 4, because the occupied and unoccupied lines contain both the school year and summer periods.

In Figure 4, the 30-day moving average of the raw efficiency scores for two buildings are shown across the year. For Building 14, the occupied and unoccupied efficiency scores were performing similar to the total score for the first 2 months. In late December, the building's efficiency score during occupied periods improves, while the efficiency score during unoccupied periods decreases. This gap remains throughout the remainder of the year. This relatively quick change in efficiency scores can indicate that an operational shift has occurred in the building, causing it to perform worse during unoccupied hours. Previous studies have highlighted how misalignment of building automation systems (Gul and Patidar 2015), and poor occupant behaviors (Masoso and Grobler 2010), can reduce building efficiency during unoccupied hours. Efficiency scores during unoccupied periods can highlight which buildings to prioritize efforts to review automation schedules and implement behavior change campaigns.

On the other hand, Building 4 performs consistently very well during unoccupied periods, with efficiency scores above 0.90. The building's occupied efficiency scores are about 10% less than the unoccupied scores for the first half of the year, with this difference increasing for the last two months of the year. The summer peak demand scores follow the same trends as the occupied periods during the summer, which is expected because their specified times during the day are very similar. For this building, it could indicate that improvements aimed at decreasing energy use during occupied hours, particularly summer peaks, may be most appropriate. In contrast to unoccupied period efficiency measures, types of efficiency measures to address occupied consumption include more capital-intensive efforts, such as retrofitting light fixtures, air conditioners, and installing demand-controlled ventilation (Koomey and Brown 2002).

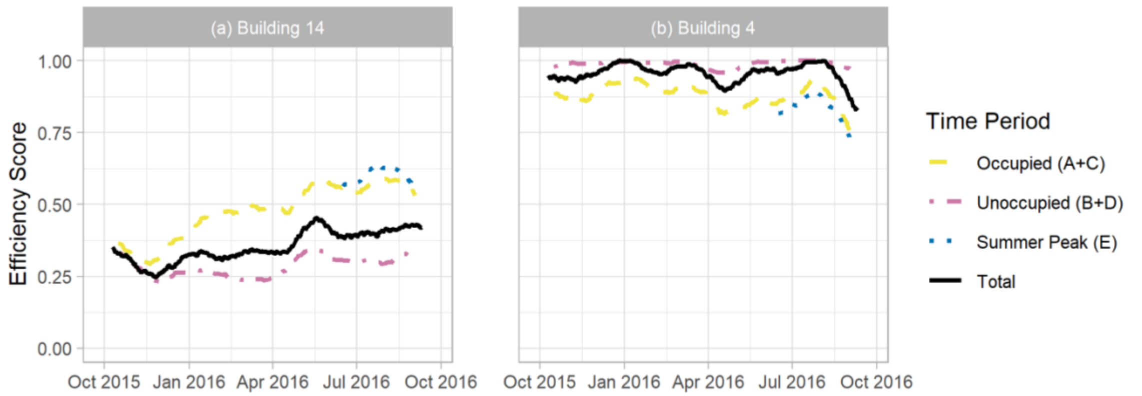


Figure 4 – 30-day Moving Average of Raw Efficiency Scores for Building 14 (a) and Building 4 (b)

2.5.3 External validation

The distribution of the regression models' fit was computed by temporal period. The means of the adjusted R-squared values for each temporal period range between 0.72 and 0.80. These values are consistent with the fit of regression models in other regression benchmarking studies (Buck and Young 2007; Chung et al. 2006; Xuchao et al. 2010). Density plots showing the distribution of the models' fit for each temporal period can be found in the Appendix (Figure 17).

The regression coefficients for each daily EUI model resulted in a distribution of coefficients for each explanatory variable. Statistically significant explanatory variables (p-values < 0.05) indicate key drivers of energy consumption across the group of buildings. Table 4 shows the frequency each explanatory variable was significant. Of the 15 explanatory variables, 14 variables significantly impacted the daily EUI for at least one day. Although all of the variables were included in energy benchmark calculations, the external validation discussion below will focus on the first seven variables in Table 4, which were frequently significant (significant for >75% of the models) for at least three of the temporal periods. The other variables were less frequently significant in the models; building age was significant for 22% to 57% of the models, depending on the temporal period, and the remaining seven variables were consistently insignificant (significant for <25% of the models), across all temporal periods.

Table 4 – Frequency of Regression Coefficient Statistical Significance

Variable	Total (n = 365)	A (n = 174)	B (n = 174)	C (n = 75)	D (n = 75)	E (n = 86)
BUR: Service	100%	100%	100%	100%	100%	100%
Floor area	99%	98%	99%	100%	100%	100%
BUR: Laboratory	94%	98%	96%	93%	91%	95%
BUR: Circulation	91%	67%	91%	72%	100%	24%
BUR: Office	88%	93%	84%	95%	91%	90%
BUR: Mechanical	85%	89%	84%	97%	88%	100%
Years since renovation	83%	62%	83%	99%	87%	65%
Building age	39%	22%	39%	57%	52%	33%
BUR: Classroom	21%	9%	15%	23%	24%	8%
BUR: Supply	18%	6%	16%	12%	32%	10%
BUR: General	11%	3%	4%	12%	21%	9%
BUR: Study	9%	2%	2%	13%	19%	5%
BUR: Special	1%	10%	1%	0%	0%	14%
Number of floors	1%	2%	1%	0%	0%	0%
Percent renovated	0%	0%	0%	0%	0%	0%

(n = number of days in temporal period)

The majority of the directions of the significant coefficients aligned with previous studies and expectations. Density plots visualizing the regression coefficient distributions for all explanatory variables can be referred to in the Appendix (Figure 18). Buildings with a higher percentage of area dedicated to laboratory or mechanical room space had higher EUIs. Such spaces have energy-intensive equipment, such as lab ventilation hoods in laboratories or IT equipment in mechanical rooms. Higher percentages of office space were also associated with higher EUIs, which aligns with findings from previous studies (Park et al. 2016). Circulation space (e.g., hallways and lobbies) was associated with lower EUIs, making sense as these spaces typically have less consistent occupancy loads and more open space without energy-intensive equipment. In addition, buildings that have not been renovated recently had higher EUIs, which can be attributed to having older, less efficient equipment such as lighting, plug loads, and mechanical systems.

For some coefficients, the relationship with is not consistently supported by the literature. Floor area was positively associated with EUI, which aligns with some studies (Park et al. 2016) and contradicts others (Chung et al. 2006). Other coefficients have yet to be examined in the context of energy benchmarking. Interestingly, service space (e.g., bathroom and janitorial areas) was positively associated with EUI. To our knowledge, previous energy benchmarking studies have not documented this variable's association with energy use. Possible drivers of energy consumption in service spaces include hot water energy use, ventilation loads, and cleaning equipment.

2.5.4 Limitations and future directions

Several limitations exist for this study, prompting avenues for future research. Common across building energy benchmarking studies, it is challenging to determine how well the benchmarking indicators agree with the actual efficiency levels of the buildings, particularly when such benchmarks are developed across large scales of buildings. In a similar vein, it is difficult to determine if the efficiency recommendations informed by temporally segmented benchmarks are the optimal efficiency improvements for the building. Future research will dig deeper into this by investigating the effect real retrofits or operational changes had on daily benchmarking results. This analysis will compare the computed daily benchmark with an associated operational or capital change and examine if the daily benchmark trends follow the expected pattern based on the particular retrofit install or operational change made.

In addition, regression-based benchmarking techniques, including the methodology developed in this study, assume the regression residuals reflect only building inefficiencies, while in reality they contain statistical noise, measurement error, and unexplained factors (Chung 2011). However, other benchmarking methodologies inherently have other limitations, such as large sensitivities to outliers and loss of physical meaning (Borgstein et al. 2016). The aim of this study was to apply existing benchmarking techniques to assess deviations between energy benchmarks during different time periods, and we opted to apply regression-based techniques due to their high interpretability (e.g., the results have physical meaning as it relates to the building) and common adoption in industry applications, such as the ENERGY STAR® score. Future studies could apply other benchmarking techniques to assess the consistency of the results across different methodologies. This is particularly relevant for benchmarking analyses performed at different scales, such as daily, weekly, or monthly, to assess the robustness and sensitivity of different techniques. Of note, building occupancy states were estimated, as measured data was not available for all buildings. Thus, estimated occupancy states may not reflect the actual occupancy levels in the buildings. Incorporating data that contains or represents a proxy for actual occupancy within each building into the benchmarking models will enhance the accuracy of the models.

In its current state, our digital twin-enabled energy management system (Figure 2) demonstrates a proof-of-concept for the platform. Future work will more deeply integrate smart meter data, temporally segmented energy benchmarks, and other resource data such as gas, heating, cooling, and water consumption. This effort will also involve user interface updates and testing with specific user groups (e.g., facility managers) to assess the utility and future direction of the platform.

2.6 Conclusion

Approaches for assessing building energy performance diverge in their ability to handle large scale analyses while still providing specific, actionable findings. Energy benchmarking methodologies can be applied across a large number of buildings, however they provide narrow insights and are limited in their ability to identify specific areas for efficiency improvement (Borgstein et al. 2016). Conversely, energy diagnosis methodologies provide more actionable energy conservation measures, however, they are most appropriate at a single-building level and require extensive and accurate data collection (Borgstein et al. 2016). This paper expands recent top-down, data-driven approaches to building performance assessments (ElYamany et al. 2017; Francisco et al. 2018a; Roth and Jain 2018) by creating temporally segmented, daily building energy benchmarks and evaluating their statistical deviations between conventional energy benchmarks.

Our findings demonstrate that across all of the buildings in the sample, temporally segmented energy efficiency scores were statistically distinct from efficiency scores during the total period for at least one temporal period. For the vast majority of buildings, such scores were statistically distinct for at least four of the five temporal periods. This indicates that for most buildings, although they may rank as efficient overall, they are not necessarily efficient during certain time periods, and vice versa. Thus, total efficiency scores mask underlying periods of inefficiencies or efficient performance. In addition, we establish that efficiency scores not only fluctuate between time periods, but also within time periods. This understanding is crucial for near real-time operational decision-making and management. Fluctuations in energy efficiency throughout the year indicate whether a building is consistently performing well, under-performing, or if a sudden change in performance occurred. This is a crucial distinction that can support decision makers develop strategies whether to investigate operational procedure modifications or opportunities for more capital intensive investments.

Overall, these results expand the usability and accuracy of traditional building energy benchmarking approaches. Temporally segmented daily efficiency metrics integrated into digital twin-enabled energy management platforms can transform approaches to energy management across a portfolio of buildings. This is of critical importance as cities are working under limited budgets to make substantial reductions in building energy consumption and strive towards 'smarter' operations. Temporally segmented building energy benchmarks show new insights using building benchmarking techniques to enable more systematic, real-time, and accurate management of city-scale building energy consumption and help urban areas reach low-carbon energy goals.

2.7 Acknowledgements

This material is based upon work supported by the National Science Foundation under Grants No. DGE-1650044 and No. CPS-1837021. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors also wish to thank Jessica Rose and others at the Georgia Tech Facilities Management Office for data access and other input.

CHAPTER 3. UNDERSTANDING CITIZEN PERSPECTIVES ON OPEN URBAN ENERGY DATA THROUGH THE DEVELOPMENT AND TESTING OF A COMMUNITY ENERGY FEEDBACK SYSTEM[2]

3.1 Introduction

Cities around the globe are heavily investing in becoming ‘smart’. World-wide investments in technology for smart city initiatives are expected to grow from more than $81 billion in 2018 to $158 billion in 2022 (Shirer and Da Rold 2018). While smart city definitions encompass broad variations, common across all perspectives is the aim to address an urban issue—whether it be energy, safety, health, mobility, or financial in nature—with an approach that is mediated through technology (Chourabi et al. 2011). Fostering citizen learning, participation, and benefit is an important element of smart city frameworks, and critiques have pointed out that this is the element most often glazed over, with greater emphasis put on technological progress and development in digitizing our cities. One prominent smart city paradigm area with a high potential to involve citizens is open data initiatives. Providing open data to citizens is envisioned to enable transparency of government operations, provide social and commercial value, and increase participatory governance (Attard et al. 2015).

[2] This chapter was published as a journal article in the Journal of Applied Energy with John E. Taylor as the co-author. The citation for the journal article is as follows: Francisco, A., and Taylor, J. E. (2019). “Understanding citizen perspectives on open urban energy data through the development and testing of a community energy feedback system.” *Applied Energy*, 256.

Within the energy sector, open data is an emerging resource made possible through technological innovations, such smart metering infrastructure (EIA 2018), and new policies, such as building energy disclosure ordinances (Palmer and Walls 2015). While individual building owners have long leveraged their own building data to improve energy efficiency decisions across residential (Zhou and Yang 2016), commercial (Gulbinas et al. 2015) , and industrial (D'Aniello et al. 2018) contexts, this new form of publicly available building data has potential to transform decision making in the context of building energy performance across entire cities. Although researchers have documented the potential value of this data to private stakeholders such as building owners, investors, and utilities, little research attention has been drawn to the use of this data to empower and promote engagement from citizens with our energy systems (Kontokosta 2013). Connecting this data with citizen interest and decision making will require transforming it in a way that improves its usability and accessibility. Energy-cyber-physical systems (e-CPSs), applied in a citizen-centric manner, present an excellent opportunity to help transform energy data to be more useful to citizens. Energy-cyber-physical systems aim to link the physical world (i.e., where citizens are most familiar) with the virtual world (i.e., where data is collected and analyzed), to enable more informed decision making (Gupta et al. 2011; Poovendran 2010). In this paper, we build on the concept of e-CPSs with a focus on citizens as the end-user in the context of what we define as *open urban energy data*. We transform this data into easily usable information through the creation of an open urban energy data feedback platform and assess citizen interactions with this system to evaluate its viability and value in supporting citizen decision making.

3.2 Background

3.2.1 Open urban energy data and the role of citizens

Public reporting of energy data through mandated building energy disclosure laws is driving a shift in the transparency of building energy production and consumption information across cities (Palmer and Walls 2016). For example, as of February 2019, 28 cities and 3 states in the US have enacted building benchmarking or disclosure ordinances (Institute for Market Transformation 2019), enabling public access to what we define as *open urban energy data*. Outside of the US, open urban energy data initiatives such as the European Union's Energy Performance of Buildings Directive have been enacted, requiring public buildings to present Display Energy Certificates (EuropeanCommission 2019). Researchers from economics domains have focused on the potential of open urban energy data to transform energy efficiency markets by supporting building portfolio owners with performance management, guiding investors with energy financing decisions, and increasing the value and marketability of commercial buildings (Palmer and Walls 2016). Noticeably, use cases and potential benefits of this data has focused on stakeholders interacting close to the real estate industry (Kontokosta 2013; Zullo et al. 2016). While the general public has been briefly mentioned as a potential user (Kontokosta 2013; Zullo et al. 2016) little research has examined in detail the potential of citizens as data users in the context of open urban energy data. As one of the core tenets of open data is to provide use and benefit to the public (Attard et al. 2015), it is worthwhile to examine public understanding of open urban energy data and their interest in using this information to help decision making.

Importantly, the release of this data coincides with a growing interest from researchers and governments of the role of citizens with our future energy systems. A substantial body of research has called for a reconceptualization of ordinary citizens from passive energy consumers to active stakeholders and innovators in creating new and more sustainable future energy systems (Bomberg and McEwen 2012; Schot et al. 2016). Citizen participation during energy project assessments and development is integral to see a project successfully come into fruition and integrate into a community (O'Dwyer et al. 2019)—citizens increase the dissemination and adoption of energy technologies (Ornetzeder and Rohracher 2006), improve the acceptance of projects or technologies (Schot et al. 2016), help designers incorporate social and environmental contexts into a project (Seyfang et al. 2010), and enhance the design of the project or technology itself (Ornetzeder and Rohracher 2013). The many ways citizens can improve energy systems reflect the broad roles citizens could have in the future in relation to energy in their community.

Expanded citizen roles with energy systems in the future will undoubtedly be mediated by technologies and data. E-CPSs integrating open urban energy data have immense potential to be a resource to citizens, and support education and decision making in their expanding roles. Integral to this effort is determining how to shape and present this data in a way that is meaningful, usable, and engaging to citizens.

3.2.2 Providing open urban energy feedback to citizens

Few researchers have focused on the best ways to shape and communicate open urban energy data to citizens. Kontokosta and Tull (Kontokosta and Tull 2015) commented on how the current format of such data is most often provided in tabular spreadsheet format, making it cumbersome to analyze and relatively inaccessible to most potential users. To address this, they created an interactive web-based platform visualizing New York City's building energy benchmarking data. While the tool was built for a variety of stakeholders, it was primarily geared towards building portfolio owners and managers. Since the development of this tool, other cities have released web-based visualizations of open urban energy data such as Chicago ("Chicago Energy Benchmarking" 2018), Los Angeles ("LA Energy Atlas" 2018), and Seattle ("Seattle Energy Benchmarking" 2018), and researchers have created platforms displaying detailed building sustainability metrics across groups of buildings (Koo et al. 2015). These tools apply a variety of techniques to communicate building energy information. While there has been a growing interest by governments and researchers in creating these platforms to improve accessibility and awareness of open urban energy data (Gulbinas and Jain 2016; IMT 2015), two critical elements have yet to be examined by researchers in the development of these platforms: (a) how to design these systems for *citizen* understanding and engagement, and (b) incorporation of user feedback during the platform *design* stage.

Expanding on the first critical element, one well-developed area of research focusing on developing tools to assist in communicating of energy data to lay audiences is energy feedback system research. Energy feedback entails communicating building energy consumption information to occupants through typically computerized means in a way that aims to be appealing to building occupants, improve awareness, and motivate pro-environmental behaviors (Fischer 2008; Sanguinetti et al. 2018). More recent literature on energy feedback has called for the expansion of energy feedback to the community-scale (Geelen et al. 2013; Pierce and Paulos 2012a). While the definition and scope of community-scale energy feedback has not yet been formally characterized, several recent studies have explored the potential of community-scale aspects within energy feedback. Burchell et al. (Burchell et al. 2016) deployed an energy feedback system as part of a larger community-scale program, and focused on how community-specific communications affected engagement. Their findings highlighted how communication strategies that are specific to a community's context, such as placing an individual's performance in the context of the community's energy events or goals, can improve the engagement with an energy feedback system. Another recent study took a different approach by creating a spatial map to communicate neighborhood-level energy consumption, and found it was useful for community groups engaging with energy issues in their community (Gupta et al. 2017). While both of these studies represent novel approaches for expanding energy feedback to the community-scale, to our knowledge an approach for designing community-scale energy feedback in the context of open urban energy data and citizen engagement has yet to be explored.

In addition, gathering feedback from users beyond the system designers and during the design stage is integral to ensure the system is appropriate for citizens as end users. User feedback can also help gather information on ways people envision open urban energy data could be useful. Literature has long drawn attention to importance of end user engagement during the design of technologies that humans will interact with (Peacock et al. 2017; Skjølsvold et al. 2017). In the context of smart energy feedback technologies, such systems have been typically designed in a way that reflects the needs and desires of the designer(s) rather than the intended users of the system (Skjølsvold and Lindkvist 2015; Strengers 2014). Geelen et al. (Geelen et al. 2013) evaluated the extent that energy technologies for smart building systems empower and enable citizens to undertake more involved and educated roles. Their findings illuminate an often observed disconnect between smart energy technology design and the end-user; design decisions rarely involved the end-user, and instead focused on technical and financial incentives grounded on the assumption of a rational end-user. Neglecting user involvement during the design phase impacts technology acceptance, engagement, and effectiveness over the long term (Abras et al. 2004; Peacock et al. 2017).

Across cities, open urban energy data is becoming increasingly available online (Institute for Market Transformation 2019; Kontokosta 2013), however, public availability of this data does not imply usability, particularly for ordinary citizens. As there is considerable research attention stressing the need for increased citizen engagement and understanding of our future energy systems (Bomberg and McEwen 2012; Schot et al. 2016), access to and interaction with open urban energy data may have a role in expanding citizen understanding and involvement. However, currently little research has focused on development of feedback tools for citizen engagement and their perspectives on open urban energy data. There are two primary objectives of this paper to help address this. The first objective is to develop a community-scale energy feedback system designed for citizens as data users and document the design and development framework for this system. In this process, we leverage open urban energy data and integrate elements from cyber-physical systems to connect virtually stored data with the built environment (Poovendran 2010). The second objective is to engage prospective users in the design of the developed system, gather feedback from expected future users about the system and assess more broadly the potential of these data for use by the general public to support decision making. It is envisioned that the developed system has potential to connect citizens with open urban energy data in a way that was previously not possible without this approach (schematic in Figure 5). In the following section, the approaches applied in the development and user evaluation of this system are described.

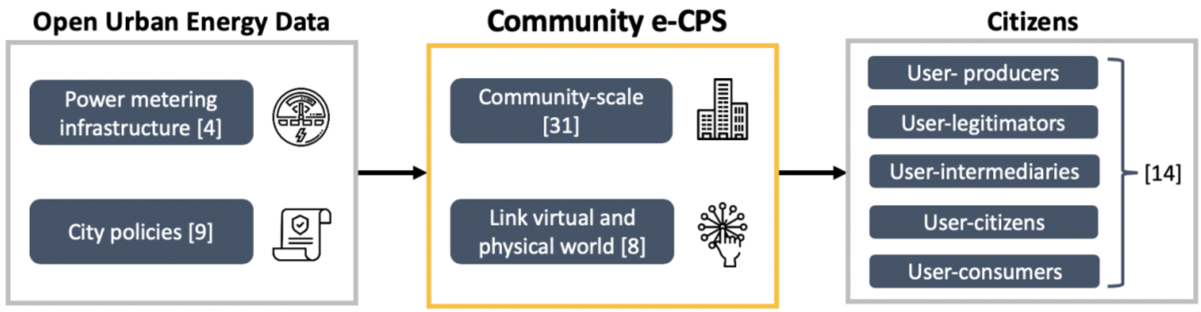


Figure 5 – High-level schematic depicting a community e-CPS as the mechanism linking citizens to open urban energy data

3.3 Methods

To carry out the first objective of the study, development and documentation of the design of a community-scale energy feedback system, we adopted a theory-driven approach (Petkov et al. 2011). This entails basing our design on findings previously established within the energy feedback literature, which is appropriate given the amount and depth of research already conducted on energy feedback system design and energy information communication. From these findings, which are summarized in the following sections, we identified three main functionalities to include in our community-scale energy feedback system: *augmented reality* feedback, *energy supply* feedback, and *energy consumption* feedback. Following the development of this system, we pursued the second objective of the study: to engage real, prospective users and gather their feedback regarding the design and information included in the system. For this aim, we employed a user-centered approach employed in the human-computer interaction field, which has been used for decades to evaluate user interfaces and gather feedback from prospective users (Abras et al. 2004). We sought to examine not only if users could accurately interpret the energy feedback system, but also how interested people are in having access to this information in the first place. Through this approach, we sought to answer the following research questions:

1. Do users accurately understand and interpret open urban energy data portrayed through community-scale: (a) Augmented Reality feedback, (b) Energy Supply feedback, and (c) Energy Consumption feedback?
2. Do people want to seek out open urban energy data provided by community-scale energy feedback interfaces? Why or why not?

The following sections detail the technical architecture, design, and user testing approach; Section 3.1 pertains to the first objective of the study, while Section 3.2 details methods related to the second objective.

3.3.1 System design and architecture

Georgia Institute of Technology's (GT's) campus served as the testbed community to build our platform on. While city functions are generally more diverse and broad compared the functions across an academic university, university campuses have been regarded has embodying the heterogeneous facilities across a community or small town, as they are composed of a combination of offices, laboratories, recreation, health, food, retail, and classroom facilities (Klein-Banai and Theis 2011). In addition, similar to city municipalities, universities manage large and diverse property portfolios, and largely benefit from energy efficiency investments over the long-term. On GT's campus, each building is equipped with a smart meter, recording building electricity consumption and production data every 15-minutes. The system was designed to communicate both electricity consumption and production data. This data is comparable to open urban energy data, which is typically released at the building level and offers time granularity of varying degrees (from annually to every 15 minutes) (BuildSmart DC 2017). The intent of the application is to place individual building performance in the context of the entire community's energy production and consumption goals as a whole. The application was built for mobile devices. The choice to create the platform for mobile device access was primarily to support one of the application functionalities (i.e., augmented reality). Additionally, mobile access to energy information also has potential to enable more timely feedback (Weiss et al. 2009) and encourage simultaneous learning with multiple users (Valkanova et al. 2013). As mentioned, a theory-based approach was adopted for the initial design of the system, as a substantial amount of empirically-tested design information exists within energy feedback. These findings are summarized in the following section, and are an expansion of the work conducted in (Francisco and Taylor 2019b).

3.3.1.1 Augmented reality feedback

Fundamentally, energy feedback systems aim to change people's energy behaviors through increased visibility and access to building energy consumption information. The Spatial view was developed to promote the visibility of energy consumption using mobile Augmented Reality (mobile-AR) (Wu et al. 2013). While application of AR technologies is limited within energy feedback research, educational fields have shown AR can increase learning by enabling users to explore technical information that is invisible in the real-world and engage with the spatial relationships contained in this information. Combining AR with portable and location-aware devices, such as mobile phones and tablets, mobile-AR affords greater context-aware and social learning through merging experiences in the real world with interactive virtual information (Squire and Klopfer 2007). Unlike conventional energy feedback displaying graphs, charts, or images on a screen, mobile-AR energy feedback combines the affordances of a real environment with the virtual world, adding more connections to a user's sense of reality. Applying mobile-AR energy feedback across buildings within a community may encourage interest in energy data, social learning, and discovery of more context-aware insights.

In the Spatial view, users look through a mobile device to visualize virtualized building energy efficiency information, demonstrated through color-coded icons, augmented on top of the physical buildings in the real world (Figure 6). As users stand or move around outside, icons appear based on their orientation and distance from a building. The color of the icon indicates the level of energy efficiency, based on a 5-color scale defined in the legend. Previous studies have encountered some user difficulty in interpreting color scales (Francisco et al. 2018b), thus the legend labels are added to explicitly indicate how a building should be performing (i.e., injunctive norm). Viewing energy use in this manner enables users to compare energy performance relative to other buildings in their community. Previous studies have shown normative feedback in energy feedback systems can evoke pro-environmental behaviors (Dixon et al. 2015; Nolan et al. 2008). The color-coded icons also display energy efficiency levels in a numerical format, as combining color-coded aesthetics with numerical representation has been found to be preferable to users (Bonino et al. 2012; Francisco et al. 2018b). To calculate a building's energy efficiency level, its Energy Use Intensity (EUI) is compared to the rest of the buildings in the sample for each time range (i.e., now, month, year). Equation 3 expresses the calculation performed to construct the yearly EUI, as follows:

$$Yearly\ EUI_i = \frac{\sum_{j=1}^{35,040} e_j}{area_i} \tag{3}$$

where i is each building in the sample, e is the electricity consumption (kWh), and j is each 15-minute increment, totaling 35,040 increments per year. Building EUI is calculated for each time range, where j is adjusted to be the number of 15-minute increments in each time interval. In addition to the augmented icons, user interactivity, which has been shown to improve engagement (Fischer 2008), is enabled through: (a) toggling efficiency values between different time ranges (buttons at top of screen), and (b) adjusting the distance from the user that the icons will appear (slider at left of screen). 'Now' indicates the building's current EUI (i.e., EUI during the most recently reported 15-minute interval), while the 'month' and 'year' are equivalent to the calculated EUI over the past 30 or 365 days, respectively. Visualization of efficiency icons have duel-coding in the map at the bottom of the screen, showing a bird's eye view of the icon locations.

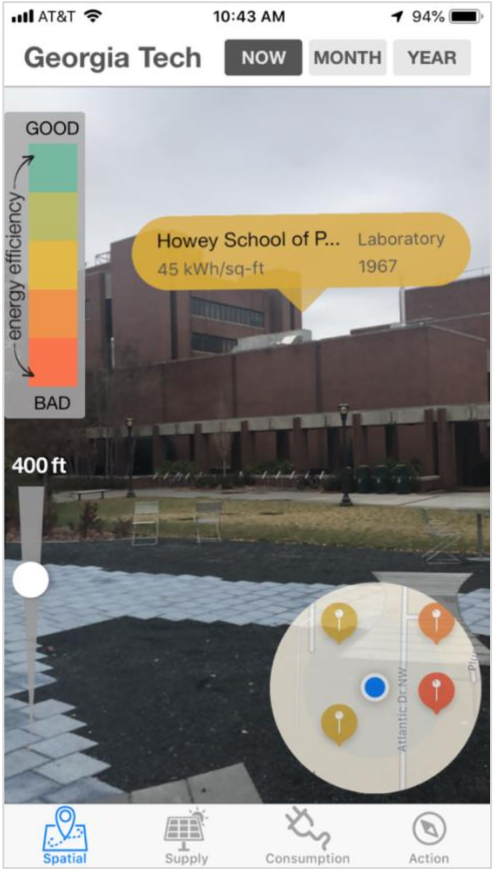


Figure 6 – Screenshot of Spatial View

3.3.1.2 Energy consumption feedback

Energy consumption is a fundamental metric represented in energy feedback systems. Energy consumption data is often transformed into a simplified metric or placed within the context of a benchmark or norm to improve the interpretability and understandability of this information (Sorrentino et al. 2019). The Energy Consumption view aims to communicate energy consumption relative to four different benchmarks: historic consumption, peer consumption, community consumption, and consumption goals. Previous studies have shown that contextualizing energy consumption can promote user understanding of performance and encourage behavior change (Jain et al. 2012). Other studies have gathered through user interviews that users can mistrust some comparisons, particularly with their peers, due to inherent differences between building characteristics and operations (Petkov et al. 2011). Combining multiple points of comparison, such as peer and historic or peer and goals may help ensure metrics resonate with users.

At the top of the Energy Consumption view (Figure 7), the campus' energy reduction goal is stated. Below, an interactive graph is provided where the user can toggle between different buildings on campus. The bar graph demonstrates the four points of normative comparison. The x-axis shows historic energy consumption from previous years. The blue bars show campus energy consumption in relation to the selected building (red bars). The horizontal lines indicate the baseline threshold from 2013 and the energy goal reduction for 2020. An injunctive norm is constructed through the emoticon on the screen, which changes based on the individual building performance. If the building's energy use increased over time, the emoticon changes to an unhappy expression. If the building reduced energy use, but at a slower rate than the rest of the campus, the emoticon expression is neutral. Injunctive norms have been found to be particularly effective at encouraging efficient performers at sustaining efficiency levels (Schultz et al. 2007). Finally, community-based language (e.g., "together we can…", "spread the word…") is included at several points in the interface to encourage a shared group identity (Burchell et al. 2016).

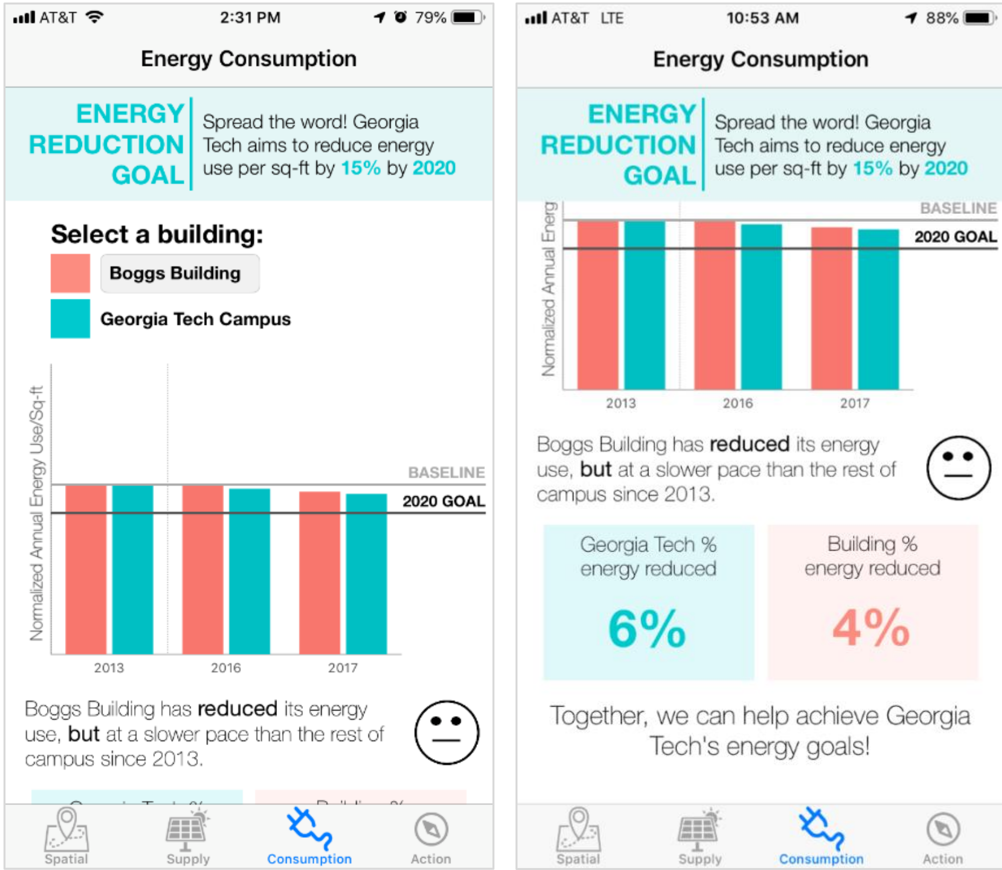


Figure 7 – Screenshot of Consumption View

3.3.1.3 <u>Energy supply feedback</u>

While traditional energy feedback platforms have focused on energy consumption, greenhouse gas emissions of energy sources delivered to a community varies drastically. In addition to improving energy efficiency, transitioning to low carbon energy sources is just as important a way to achieve clean energy goals. Studies have recognized that transitioning to low carbon sources will likely involve greater citizen engagement, such as coordinating with neighbors to help match supply and demand for individual or co-generation (Geelen et al. 2013). The Energy Supply view aims to connect users not only with how energy is consumed in a community, but also how it is supplied.

The Energy Supply view displays a breakdown of electricity sources, from fossil fuels to renewables, in the context of the community's renewable energy goals (Figure 8). The header of the screen contains the renewable energy target for a community. Goal setting has been implemented in many energy feedback applications (Gulbinas et al. 2014) and has been shown to motivate pro-environmental behaviors (Abrahamse et al. 2005; Wood and Newborough 2007). Below the goal, an interactive chart displays the community's progress in reaching their goal. The types of energy sources provided include coal, nature gas, nuclear, and renewable energy sources. In this case, the percentage breakdown for each source came from Georgia Power, the sole electricity provider to the campus. Thus, building comparisons of electricity sources only change if a building produces its own electricity, such as through a solar installation or the use of geothermal. Users can interact with the graph to compare any building's electricity source breakdown to that of the entire campus in the context of the campus's renewable energy goals. Normative comparisons are complemented by injunctive norm, similar to the Energy Consumption view. Normative feedback literature has documented a boomerang effect, where users that are performing better than average sometimes reduce their performance. Injunctive norms have been found to help combat this behavior by indicating the approval of their actions (Schultz et al. 2007).

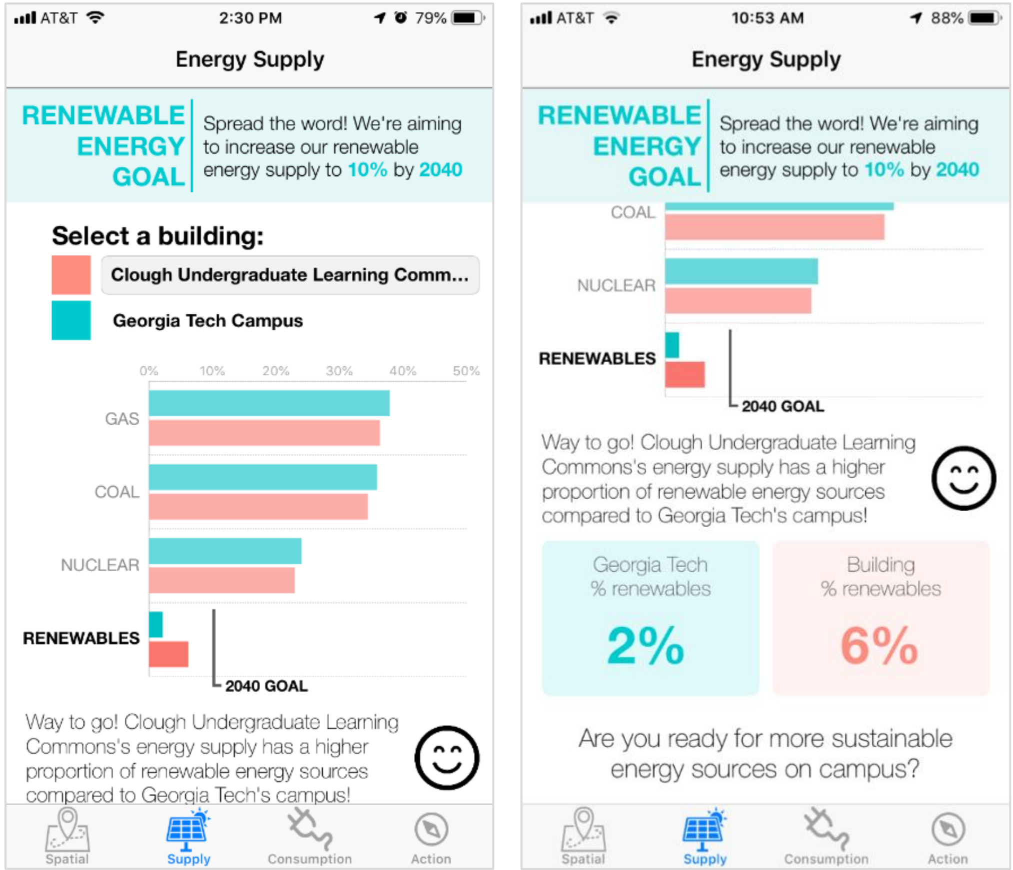


Figure 8 – Screenshot of Supply View

3.3.2 Platform architecture

The mobile-based application was created by the authors. We chose to develop the application using Swift 4 programming language and the Xcode Integrated Development Environment (IDE), thus the application is available on iOS mobile devices. Multiple packages were used to support the development. Within the Spatial view, ARKit, MapKit, CoreLocation and SpriteKit frameworks (“Apple Developer” 2018) were applied to implement augmented reality effects, render graphics, estimate geolocation and orientation, and generate maps. User location was updated every 0.5 seconds (extracted by the *userLocation* instance property in the class *MKMapView*), providing the user with their exact location in the bottom-right map. Using the user's current location, a function calculated all buildings within the user-specified radius (default being 300 feet), which dictated which building icons and information appeared to the user. Back-end implementation of the AR capabilities relied extensively on the open-source code provided by ProjectDent (AndrewProjDent 2018). Generation of graphs displayed in the Energy Supply and Consumption views was designed and implemented using the Charts framework (Gindi and Jahoda 2016). Within each view, multiple IBOutlets and IBActions facilitated user interaction (e.g., pressing the time range button, selecting a building from a dropdown). Electricity and building data were cleaned and summarized using the R programming language, then compiled into .json format, and stored directly within the application. In future iterations of the application, data will be retrieved by the application through an API to enable near real-time (15-minute interval) energy updates.

3.3.3 Citizen-centered evaluation

Drawing from user-centered design best practices within human-computer interaction fields (Abras et al. 2004), the second objective of the paper involved a pilot study to evaluate the community-scale energy feedback system developed above from the perspective of prospective citizen users. A wide range of quantitative and qualitative methods have been employed to gather data on user perceptions and understanding of technologies. Quantitative methods are generally used to test products where the user needs are already well-defined. As community-scale energy feedback is a relatively novel concept to most people, the aim of our user testing was exploratory and prioritized collection of qualitative data.

For the evaluation, 16 study participants were recruited to complete two activities: (1) a thinking aloud session, and (2) a survey questionnaire. Thinking aloud methods have been used for decades by researchers to diagnose usability issues and improve user interfaces (Nielsen 1994). The thinking aloud procedure involved a one-on-one session between a researcher and participant, where the participant was instructed to vocalize their thoughts out loud as they interacted with the interface. To prompt user interactions, they were provided with tasks to direct the user to test specific functionalities. Before beginning the thinking aloud session, the researcher provided participants with details on how to 'think out loud', such as specifying what they believe is happening, why they are taking an action, and what they are trying to do. The same researcher tested each participant and read the same instructional script prior to the session. While the participant completed each task, the mobile device's screen and microphone recording were turned on. This provided data on how the participant accomplished each task and what their quasi-raw thought stream was as they encountered different functionalities.

Immediately following the thinking aloud session, the participants completed a web-based survey questionnaire. The survey was divided into three parts. First, participants were asked questions to assess how accurately they interpreted each functionality. We opted for open-ended responses to these questions, as multiple choice options could influence how a user reports their understanding of the interface. Because this is an exploratory pilot study without statistical interpretation, we argue it is best to capture raw user thoughts rather than pre-constructed, limited responses to assess user interpretation. A list of these questions can be referred to in Table 5. The second part inquired more broadly about users' opinions on community-scale energy feedback. It captured users' desire to have access to this information, at what geographic scale, as well as motivations for wanting to seek out this information. A full list of questions and response types can be referred to in Table 6. The third part of the survey included demographic questions to learn more about respondents. Sixteen participants were recruited to take part in the user testing. While we expect that the primary users of this system would be people who are professionally or personally interested in local energy issues, we recruited people from a variety of backgrounds, from inexperienced to energy experts. This was done to gather a more holistic understanding of user interpretations. The user testing was approved through IRB protocol #H18398 and participants were compensated monetarily for taking part in the study.

Table 5 – Participant Interpretation of Community-scale Energy Feedback Features (Survey Questionnaire Part 1)[3]

Question ID	Question	Question Type
AR1	If a building's icon is red, what do you think this indicates about the energy efficiency of the building?	Open-ended
AR2	If the button "Month" is selected, why do you think the color of some icons change?	Open-ended
S1	What do you think the happy face on the screen indicates?	Open-ended
S2	Based on the screenshot above, how much of the Georgia Tech campus energy supply comes from renewable energy sources?	Open-ended
S3	Based on the screenshot image above, has the Georgia Tech campus reached its renewable energy goals yet?	Open-ended
S4	Based on the screenshot image above, most of Georgia Tech's energy comes from what resource?	Open-ended
C1	What do you think the 'Baseline' label refers to?	Open-ended
C2	What building(s) do the red bars in the graph refer to?	Open-ended
C3	Based on the screenshot image above, has the Boggs building reduced its energy use at a faster or slower pace compared to the rest of the campus?	Open-ended
C4	Based on the screenshot image above has Georgia Tech reached its campus energy consumption goals yet?	Open-ended
Overall	As a whole, do you think this application offered: (A) Too much information, (B) Too little information, and (C) Just the right amount of information. • (*conditional follow-up for responses A or B*) Why? If you could add/remove one feature, what would it be?	Multiple Choice

[3] Each of these questions were followed with a 5-point Likert scale question asking users to report how confident they were in their answer, which are not listed in the table above for brevity.

Table 6 – Participant Desire and Motivations for Seeking out Community-scale Energy Feedback (Survey Questionnaire Part 2)

Question ID	Question	Question Type
B1	In general, how interested are you in having access to a Community Energy Feedback System in the following locations: - Georgia Tech Campus. - The neighborhood or community I live in. - The city I live in.	5-point Likert Scale
B2	Please specify how likely or unlikely you think a Community Energy Feedback System would motivate you to: see list of behaviors in Table 9.	5-point Likert Scale
B3	How often do you think you would seek out the information provided by a Community Energy Feedback System? (A) Daily (B) Weekly (C) Monthly (D) A few times (E) Once (F) Never (G) Only for specific occasions	Multiple Choice
B4	We are interested in why people would want to seek out information included in a Community Energy Feedback System. Below, please describe why you would want to have access to such a system. If you do not want to have access, please describe why.	Open-ended

3.4 Results

Sample screenshots from the resulting community-scale energy feedback system can be referred to in Figure 6, 7, and 8, and it source code with additional documentation of its capabilities is available for download within a Github repository. Thinking aloud and survey data was collected from December 2018 through January 2019. A total of 16 participants completed the user testing activities. As an important aspect of the developed application is that users are familiar with the community it represents, participants were required to identify as a member of the GT community. Participants' affiliation with the GT community consisted of being an undergraduate student (n=6), Master's student (n=4), PhD student (n=4), or staff member (n=2). About half (n=9) of the participants identified as female and the remaining participants identified as male. The survey also inquired about people's previous familiarity with AR technology and their previous involvement and interest in energy issues. These results are presented in Figure 9 and 10. In the following sections, the results pertaining to each of the research questions will be discussed.

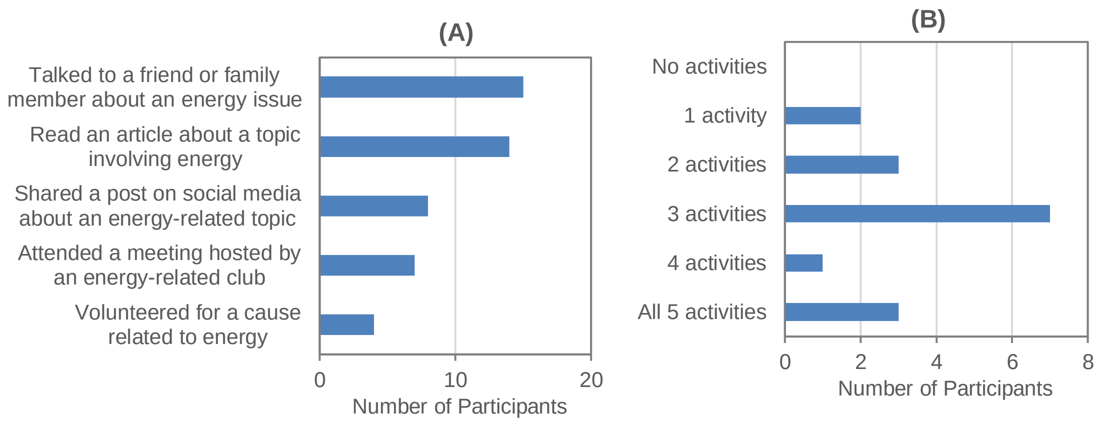


Figure 9 – (A) Results from the survey question, "Consider the following and check all that apply. In the last three months have you:"; (B) Total number of activities each participant reported from chart (A)

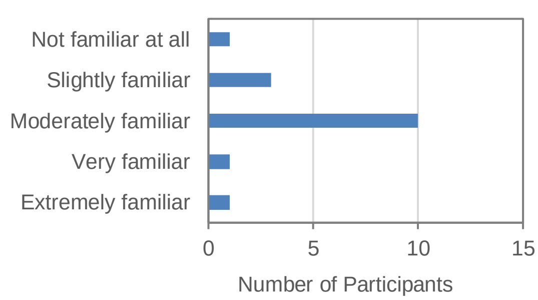


Figure 10 – Results from the survey question, "Prior to this study, how familiar were you with Augmented Reality technology?"

3.4.1 Do users accurately understand and interpret community-scale: (a) Augmented Reality, (b) Energy Supply, and (c) Energy Consumption feedback?

Participants' ability to accurately interpret the interface was assessed through the thinking aloud procedure and the survey questions listed in Table 5. A summary of the number of participants who answered each accuracy question correctly and their associated confidence in their response is provided in Table 7. Importantly, the thinking aloud procedure shed light on how a participant's understanding of each feature evolved as they completed each task. Understanding tended to improve as their exposure to the interface increased. As a result, sometimes participants initially misinterpreted a particular feature in the thinking aloud portion, while at the same time they were able to accurately interpret the same feature in the follow-up survey. The combination of these methods enabled a more in depth understanding of the user learning process. Such events are explained in more detail in the following paragraphs, which discuss each of the application's features.

Table 7 – Count of participant confidence levels for each accuracy question assessing accuracy of interpretation.

Question ID	Number of Correct Responses	Very confident	Confident	Slightly confident	Not confident
AR1	15	14	2	0	0
AR2	15	5	8	2	1
S1	13	9	7	0	0
S2	16	13	2	1	0
S3	15	15	1	0	0
S4	15	14	2	0	0
C1	11	7	6	2	1
C2	16	15	1	0	0
C3	16	13	0	3	0
C4	15	15	0	1	0

3.4.1.1 Augmented reality feedback

In the Spatial view, participants were able to use the augmented reality feedback to identify a building's energy efficiency level with relative ease and accuracy. When asked in the survey, "*If a building's icon is red, what do you think this indicates about the energy efficiency of the building?*" (AR1), the vast majority of participants accurately interpreted the meaning of the color scale with high levels of confidence (Table 7). This was also evident from the thinking aloud exercise, where participants were able to interpret the meaning behind the colors without difficulty or error. Furthermore, the double-coded numerical and color-coded efficiency representations also appeared effective; some participants focused initially on the colors to determine efficiency levels, while others were more inclined to focus on the numerical representations. By the end of the exercise, participants tended to use both. Notably, the building characteristics integrated into the color-coded icon (e.g., year built, building type) stimulated comments from the participants as to why they thought a building had a certain efficiency level in relation to its listed characteristics, or how they were surprised by the results. For example, after visualizing a building with a green icon, one participant commented, "*that's interesting, oh and it was built in 1988 it appears, compared to this one that was built in 1967, so that's surprising. I would think that it would be less energy efficient since it's older, but it's not*".

One oversight by the participants in the Spatial view was highlighted by the survey question, "*if the button "Month" is selected, why do you think the color of some icons change?*" (AR2). While the vast majority (n=15) answered this question correctly, people were less confident in their response compared to the other survey accuracy questions. From the thinking aloud session, it was notable that most users failed to notice the time range buttons at the top of the screen. Thus, the lower confidence levels may be reflective of the minimal interaction participants had with these buttons during the tasks. In addition, participants commented in the survey that they would prefer to have more detailed information about the time ranges. Specifically, several commented they wanted to know if the time ranges reflected averages (i.e., the annual energy use divided into monthly or daily averages) or represented real-time changes in energy use. In a similar vein, participants also were looking for more detailed information about the energy efficiency color scale. More specifically, they wanted to know what 'bad' or 'good' was in reference to (e.g., the GT campus, national averages, etc.).

3.4.1.2 Energy supply feedback

In reference to the Energy Supply feature, participants were asked, "*What do you think the happy face on the screen indicates?*" (S1). While 13 respondents answered this question correctly, they were relatively less confident in their responses. From the thinking aloud session, participants expressed confusion about what level of building performance 'deserved' a smiling face. Moreover, it was not clear what emoticon options a building could potentially achieve. One participant summed up these concerns tellingly with, "*what is the smiley face scale*?". Questions S2, S3, and S4 were all answered with high rates of accuracy and confidence, indicating that participants were accurately able to identify from the bar graph Georgia Tech campus' current level of renewable energy production, renewable energy goal, and the energy resources the campus supply is composed of. An important main design issue in the Energy Supply page was determined through the thinking aloud activity, where most participants took a long time to notice the renewable energy goal listed at the top of the screen. Instead, their eyes and attention went immediately to reading and interpreting the graph on the center of the page.

3.4.1.3 Energy consumption feedback

In reference to the Energy Consumption feature, participants were asked, "*What do you think the 'Baseline' label refers to?*" (C1). Compared to the rest of the accuracy questions, this question had the lowest correct response rate, and one of the lowest confidence rates. While most of the participants (n=11) understood that the 'baseline' represents a reference point to compare a building's current energy efficiency to, only 6 participants reported that the baseline referred to a building's performance in the year 2013.

From the thinking aloud session, most participants (n=9) had noticeable trouble interpreting the meaning of the 'baseline' and '2020 goal' horizontal lines on the Energy Consumption page. While most participants were eventually able to correctly interpret the graph, they commented that they were initially confused because on the previous Energy Supply page, a building achieves the goal when its bar exceeds the '2040 goal' line. Conversely, the way the Energy Consumption graph was designed, the farther the bars are below the goal line, the better a building is performing relative to the goal. This created confusion when the conceptual model for interpreting the graph was reversed between the Energy Supply and Consumption features (i.e., wanting to go below instead of above the goal line).

Questions C2, C3, and C4 were answered with high rates of accuracy and confidence, showing that by the time of the survey users were able to accurately and confidently interpret from the bar graph what the red bars refer to, a building's level of energy reduction compared to the campus, and the energy reduction goal. With regards to C4, which inquired, "*Based on the screenshot image above has Georgia Tech reached its campus energy consumption goals yet?*" it is important to note that while users answered this question correctly and confidently, many had trouble interpreting the graph when they were initially encountered it, as described in the previous paragraph. This demonstrates the effectiveness of the thinking aloud activity combined with the survey in helping understand the users learning process and difficulties.

3.4.1.4 Overall application feedback

Judging the application as a whole, 75% of survey respondents (n=12) reported the application provided 'just the right amount of information'. The remaining participants (n=4) indicated it provided ‘too little information’. No respondents felt the application provided ‘too much information’. A follow-up question was asked when respondents selected ‘too little’ or ‘too much’ information, inquiring about what they would like to add to or remove from the application (this was presented on the next page, after submitting the answer to the previous question). For the four respondents who were asked this follow-up question, two themes emerged. First, two commented they would like to see the information in the Spatial feature better integrated with the information in the Energy Supply and Consumption feature. As one participant explained, “*The Spatial Tab tracks energy efficiency, but that information is not available for easy searching. Conversely, the supply and consumption tabs do not offer an interesting spatial visual for their respective metrics*”. This could potentially be accomplished by allowing the user to transition to the Energy Supply or Consumption graphs by clicking on a building icon in the Augmented Reality feature. Design improvements will be discussed in more detail in the Discussion section. In addition, the second theme focused on wanting more detailed information about retrofits implemented or sustainability features for each building. Participants reported they wanted this information to understand better *why* a building may be performing poorly or efficiently.

3.4.2 Do people want to seek out the information provided in community-scale energy feedback interfaces? Why or why not?

Participant openness and desire to seek out the information provided in a community-scale energy feedback interface was assessed by the questions listed in Table 6. For the open-ended question, “*We are interested in why people would want to seek out information included in a Community Energy Feedback System. Below, please describe* *why you would want to have access to such a system. If you do not want to have access, please describe why*" (B4), several trends emerged. These trends were aggregated and sorted into three categories: (a) individual motivations, (b) motivations in relation to their peers, or (c) motivations in relation to their institutions (Table 8). In Table 8, specific motivations and the number of participants who mentioned each motivation is grouped by each category. As the question was an open-response, participants' responses could include comments related to multiple categories or motivations. The total number of participants who had at least one comment in a category is specified in the first column. The most frequent motivations were related to the 'individual' category, and are comprised of motivations driven by personal interest, values, or financial reasons. Motivations belonging to the second group were less frequent and involved commentary in relation to their peers or community. Comments falling under the last category were least common, but covered a wide range of concepts related to government or institutional structures. For example, one participant commented that access to community-scale energy feedback, "*would make me want to use the data to lobby for policy changes or programs that could help expand energy efficiency upgrades at a community level. I think the visualization is most helpful for outside my home and thinking at a neighborhood, campus, or city level*".

Table 8 – Summary of participant open-ended responses to why they would or would not want to have access to a community-scale energy feedback interface.

Category	Motivation	Count
Individual (n=12)	Financial- to save money or inform purchasing/renting decisions	8
	Curiosity - to stay informed or have fun	7
	Values- care for the environment or their community	3
Peers (n=7)	Learn about how their building is performing compared to others	5
	Learn about community buy-in to energy goals	1
	Promote peer learning about energy use	1
Institutional (n=4)	Hold cities/institutions accountable to goals and lobby for better practices	2
	Learn about energy related programs in their area	1
	Support neighborhood energy organizations with targeting efforts	1
	Support energy efficient businesses	1

A few participants commented on why they would not want access to or use a community-scale energy feedback system. One participant expressed they do not think they would check out this type of feedback system unless there was a stimulus that prompted them to be curious about building energy use (e.g., the building looked new and/or efficient, when buying a home). In addition, another respondent expressed that they felt that public reporting of residential data would be an invasion of privacy, and they would not be supportive of open access to this type of data.

Participants were also asked about how likely or unlikely they were to change certain behaviors after having access to a community-scale energy feedback system (B2). A summary of these results is provided in Figure 11. Of all of the behaviors listed, participants were most likely to report that a community-scale energy feedback system would motivate them to check a home's energy efficiency before buying or renting decisions. Fewer participants selected that this type of feedback system would encourage them to attend local energy meetings or participate in an energy related club. It is important to note that differences in the responses between the behaviors is largely dependent on the types of behaviors participants are predisposed to (i.e., in general less people are likely volunteer for a club than use something for personal benefit). The effect of community-scale energy feedback on behaviors should be compared to baseline levels of behaviors to gain an accurate understanding of its impact. However, these preliminary results speak to the wide variety of behaviors community-scale energy feedback systems could potentially impact. Across all of the behaviors mentioned, at least 10 participants selected that a community-scale energy feedback system was 'somewhat likely' or 'extremely likely' to change that associated behavior.

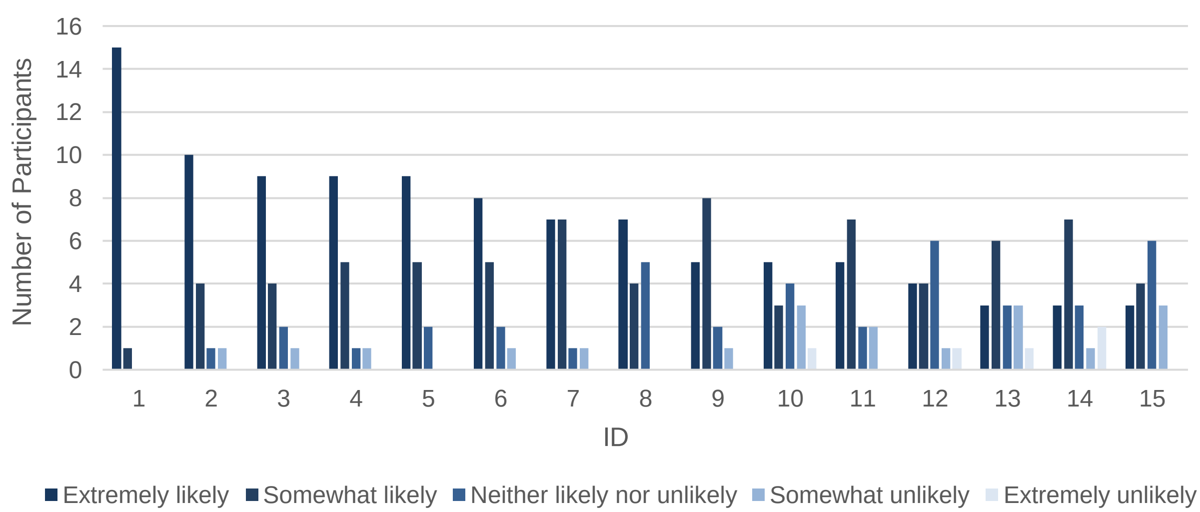


Figure 11 – Participant responses to how likely or unlikely they are to use a community-scale energy feedback system for each behavior (refer to Table 9 for the list of behaviors)

Table 9 – List of behaviors referenced in Figure 11

ID	Behavior
1	Check a home's energy efficiency before buying
2	Check a home's energy efficiency before renting
3	Buy smart technologies
4	Talk to my neighbors or peers about energy use
5	Help to try and meet energy consumption goals in my community
6	Invest in renewable energy
7	Purchase more energy efficient equipment or technologies
8	Opt to go to a place because it is efficient
9	Change energy behaviors (e.g., turn off lights or appliances more)
10	Call my representatives about energy issues
11	Become involved with an energy project in my community
12	Choose to shop at a different store because it is inefficient
13	Attend a local meeting about energy in my community
14	Share information about energy use on social media
15	Participate in a local energy club

Participants were also asked to specify how interested they were is having access to community-scale energy feedback spanning different geographic regions (B1). All locations, including the Georgia Tech campus, neighborhood they live in, and city they lived, received overwhelmingly positive responses (Table 10). When asked how often they would use such a system (B3), 5 participants responded ‘weekly’, 7 responded ‘monthly’, and the remaining 4 reported ‘a few times’ (‘daily’, ‘once’, and ‘never’ received zero votes).

Table 10 – Results from the survey question on what geographic scale the respondent is interested in having access to community-scale energy feedback

Location	Extremely interested	Somewhat interested	Neutral	Somewhat uninterested	Extremely uninterested
Georgia Tech Campus	7	9	0	0	0
Neighborhood or community I live in	10	4	1	1	0
City I live in	8	7	1	0	0

3.5 Discussion

The first aim of this study was to develop and document the design framework for a community-scale energy feedback system, a novel approach for connecting community residents with open urban energy data. Building from previous work incorporating community-based communication and mapping elements into energy feedback (Burchell et al. 2016; Gupta et al. 2017), we developed a platform facilitating user connections between building energy data and their physical community through augmented-reality visualization strategies and graphical views of energy use and supply. While visualization platforms have been developed in response to the increasing release of open urban energy data, these platforms have been primarily intended for use by entities in the real-estate sector or those with the capital and resources to make sense of the data (Gulbinas and Jain 2016; Kontokosta and Tull 2015). In response to calls for reconceptualizing citizens as active innovators and stakeholders in preparation for significant energy transitions (Bomberg and McEwen 2012; Schot et al. 2016), we document the design principles and develop a prototype community-scale energy feedback system designed for use by ordinary citizens or community members. In addition, our second aim was to evaluate the prototype by gathering feedback from prospective users. To engage prospective users, thinking aloud procedures were applied, which have been used effectively for prototype evaluation in the field of human-computer interaction for decades (Nielsen 1994). This evaluation approach was critical as engaging prospective users during the design stage is an often overlooked step in the development of energy technologies, which can substantially impact the success of these technologies during implementation (Geelen et al. 2013; Skjølsvold and Lindkvist 2015).

The thinking aloud and survey results indicated there was high interest among most participants in having access to open urban energy data across all three geographic community scales (i.e., the Georgia Tech campus, the neighborhood, and the city level). These results agree with previous community-scale energy feedback studies that established citizen interest in energy data at the neighborhood scale (Burchell et al. 2016; Gupta et al. 2017), and expand positive citizen interest to the campus (e.g., universities, workplace campuses) and city scale, which have yet to be empirically investigated by research. This provides support for the development of campus and city-scale energy feedback, which from a data availability and privacy perspective, whole building data at the campus and city-scale is more readily available due to emerging open data requirements (Zullo et al. 2016), and may be more feasible to implement compared to neighborhood-level feedback.

Regarding why participants wanted to have access to such a system, the open-ended responses had substantial variation. Aggregated responses showed participants wanted to seek out this information for individual, peer-related, or institutional-related reasons.

Individual motivations were most commonly listed (e.g., financial or ethical reasons). Nevertheless, the responses show a wide variety of ways people reflected that community-scale energy feedback could affect them and the way they engage with their energy systems. The diversity of comments corroborated previous work examining the diverse roles citizens can have when engaging with emerging energy systems (Schot et al. 2016). This has implications for literature developing frameworks for energy consumption behaviors across a group or campus of buildings (Azar and Al Ansari 2017) by informing considered behaviors in future work. This also has implications for future experimental work on community-scale energy feedback; the diversity of citizen roles in energy systems and broad scale of community-scale energy feedback behaviors widens the range of variables researchers could potentially measure to assess behavior change. An important consideration for future research, however, is designing how to measure the potential effects of this type of feedback.

In comparing the Augmented-Reality, Energy Supply, and Energy Consumption feedback, the AR view appeared to be the easiest for participants to understand. Users quickly and accurately understood the meaning behind the color-coded icons and additional building data within the icons, while they were slower to interpret the Energy Supply and Consumption views correctly. These findings agree with the results of other studies examining the impact of color-coded spatial views (Bonino et al. 2012; Francisco et al. 2018b), where users reported color-coded information helped them understand energy use and found this information more intuitive compared to typical bar charts. In addition, one of the advantages of augmented reality is it can integrate greater amounts of information into the visualization (Wu et al. 2013). In the AR view, this feature was able to integrate building characteristic data into the visualization, stimulating comments regarding efficiency levels in relation to these characteristics. Participants still found utility in the graph view information, but preferred this information to be better integrated with the AR view. This supports a previous study where people preferred color-coded and numerical data to be integrated together (Bonino et al. 2012). We extend these findings by suggesting that integration of AR features with graphical views has the potential to improve engagement and understanding by representing the information in a multitude of ways.

Another trend that emerged was participants wanting more information in general within the application. For example, participants desired more information about why a building was performing efficiently or inefficiently (e.g., sustainability features, recent renovations). Some participants even became skeptical of the data because there was not additional information to provide context for the reported efficiency levels. Providing detailed building retrofit or system information, particularly at larger scales, is a challenge as it is difficult to collect, standardize, and maintain the reporting of this information. The implication of this finding is that future design and deployments of community-scale energy feedback systems should strive to provide context to users to improve their trust in the information, while also limiting the information scope so that it is feasible to maintain and ensure the accuracy of the feedback.

Several limitations exist in this study, prompting avenues for future research. Regarding the design and development of the application interface, while our current community-scale energy feedback system has the potential to increase user understanding of community energy systems, incorporation of action-oriented elements is currently limited, which are important for behavior change (Jensen 2010). This is part due to the low granularity inherent in whole building data, which does not allow for disaggregation of energy use by behavior or appliance to give users feedback on the impact of their own behaviors. At the same time, installing more granular meters, such as plug-load monitoring, is not financially viable in a commercial context for many buildings (Wang et al. 2017). Given that open urban energy data (e.g., whole building data) is substantially more feasible to collect, maintain, and report at scale–and as such has become cities' primary public output for building energy data–we argue that research examining potential citizen use of this data has critical and meaningful value. Furthermore, granular building data is not required to add action-oriented elements; incorporation of energy tips, which can be personalized to one's own community, as well as adding interaction between users to support social engagement and connection should be considered by researchers looking to expand on community-scale energy feedback.

Regarding the system evaluation, applicability of our findings to new contexts should be carefully considered. The platform was designed for the GT campus, a university located in an urban setting, and the study's evaluation engaged people identifying as members of the GT community. As documented through the demographic data, most participants had previous involvement or interest in energy issues. Thus, our evaluation results capture responses to community-scale energy feedback from the perspective of relatively energy-cognizant populations in a university setting. Extending this system across an entire town or city will likely elicit different and more varied feedback based on the new context and less homogenous population. The results of this study provide a basis for such an extension; community-scale energy feedback is a relatively novel and promising concept, and the aim of this paper is to propose an initial proof-of-concept for others to build from, critique, and apply to new contexts. Relatedly, inherent in the creation of energy feedback systems, which are designed to reach a broader population, is that there is no design that will engage everyone; differing values, routines, and interests will impact adoption and use across populations. We expect that the initial adopters of community-energy feedback systems, for both university and city settings, will be those with existing interest or involvement with energy issues. Therefore, those who participated in this evaluation were likely representative of people most likely to adopt this system at GT.

At the same time, the aim of this application is ultimately to engage and support citizens with decision-making and action regarding energy-related behaviors. We acknowledge that, particularly when expanding such systems to a larger urban context, citizen use of data is not solely dependent on individual interest and ease of access to data, and that many other factors–such as prior knowledge, resources, and constraints–will influence a person’s ability to make effective use of this data (Gurstein 2011). Striving towards open data’s promise for empowering citizens and promoting democratic action requires purposeful collaboration with existing programs, incentives, or workshops to promote broad use of data. Many community-based energy initiative researchers have demonstrated how tools and feedback technologies can be most effective when integrated with existing community programs to promote broad citizen use, engagement, and benefit (Burchell et al. 2016; Gupta et al. 2017). This is critical for researchers to consider in the testing of the impacts of community-scale energy feedback systems.

Finally, while there were not enough participants to perform a statistical evaluation, limiting the number of respondents allowed the evaluation to include more extensive data from each participant that would not have been possible with a larger group. Capturing qualitative and descriptive data was also important given the novelty of community-scale energy feedback systems to participants; this type of data facilitates a more nuanced understanding of participant perspectives on open urban energy data and avoids biased results from prescribing perspectives. The results of this broad examination establishes a foundation to improve quantitative and qualitative evaluations in future studies.

3.6 Conclusion

As cities invest in technologies to solve urban issues and become ‘smart’, vast quantities of open data will be produced and made available for public use. In this paper, we define this data as open urban energy data. For this data to be useful, particularly to citizens, it is important for it to be shaped in a way that is accessible, engaging, and actionable. In the area of energy and sustainability, this is of critical importance as our future energy systems will require a deeper level of understanding and engagement from citizens in order to reach our sustainability aims. Energy-cyber-physical systems have immense value in working to transform data in the virtual world and link findings to our physical reality in order to improve understanding and decision making. Based in an energy-cyber-physical system perspective, our objectives for this study were to connect open urban energy data to citizens by: (a) developing a novel community-scale energy feedback system, and (b) evaluating this system using a user-centered approach. The developed system applies new visualization techniques in energy feedback (i.e., augmented reality) to enable links between physical infrastructure and smart meter data, and complements this feature with interactive, graphical displays of data. We involved 16 prospective users in the evaluation process to assess how accurately they interpreted the feedback system and gauge how interested they are in having access to this type of emerging data. The results can be used to identify specific strategies to improve the system design, indicated high interest among participants in having access to this system ($\geq$ 85% were somewhat or extremely interested) across all geographic scales presented (campus, neighborhood, and city), and established a foundation for future work in the area of open urban energy data feedback to citizens. Overall, this study presents the development and evaluation of a novel citizen-centric community-scale energy feedback system integrating open urban energy data. As open data becomes a prevalent potential resource in the era of smart cities, energy-cyber-physical systems have the potential to improve the accessibility of this information for citizen benefit. This study presents a critical step examining technologies' role in engaging the public to help achieve a sustainable, low-carbon, and people-oriented energy future.

3.7 Acknowledgements

This material is based upon work supported by the National Science Foundation under Grants No. DGE-1650044 and No. CPS-1837021. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors also wish to thank Jessica Rose and others at the Georgia Tech Facilities Management Office for data access and other input.

CHAPTER 4. DOES URBAN FORM IMPACT SOLAR PHOTOVOLTAIC SYSTEM ADOPTION? EXAMINING RESIDENTIAL SOLAR IN THE SOUTHEASTERN U.S.

4.1 Introduction

Built infrastructure systems are interdependent and long lasting. From buildings to power generation to information and communication technologies, interactions between infrastructure systems are expanding as cities seek to meet the needs of increasing populations while being pressed to become more livable, sustainable, and resilient (Chester and Allenby 2019). An emerging body of work has emphasized the need to better understand relationships between infrastructure systems because infrastructure's relative permanency affects available climate change solutions over the long term (Creutzig et al. 2016).

Within the context of urban planning, considerable literature has assessed interactions between the shape, size, and density of built infrastructure (i.e., urban form) and infrastructural or environmental outcomes. For example, studies have found urban form can impact land use arrangement outcomes (Ratti et al. 2003), energy consumption levels (Gupta 1997), and mobility patterns (Camagni et al. 2002). Less examined has been the interaction between urban form and emerging types of infrastructure systems in cities, such as solar photovoltaic (PV) systems. Building rooftops (and less frequently, facades) are critical infrastructure supporting an increasing number of solar photovoltaic (PV) systems in urban environments. The US EIA's 2020 Energy Outlook expects the US to radically shift its electricity mix by 2050, predicting solar to increase from serving 15% of the US renewable electricity generation to 46% (U.S. EIA 2020). At the same time, cities are densifying building infrastructure to accommodate more people, increase efficiency in services delivered, and reduce greenhouse gas emissions. To date however, evaluation of how varying urban forms may restrict or potentially encourage solar adoption remains largely unexamined in research.

This paper examines if and how urban form relates to solar PV adoption across urban regions in the Southeastern US. City urbanization practices and efforts to become more resilient and sustainable will result in city infrastructure undergoing substantial changes in the coming years. As solar PV deployments are built upon part of this infrastructure (i.e., rooftops), it is important to examine and quantify the potential dependencies between these two systems. Developing cities in a way that is not cognizant of infrastructure dependencies between solar PV and the built environment runs the risk of investing in long-lasting infrastructure systems that limit the viability of solar PV or are economically cumbersome to modify in the future. Greater understanding of the relationship between these two systems can proactively inform urban planning decisions that have impacts over the long term to increase opportunities for climate solution pathways.

4.2 Background

Within solar PV adoption literature, determinants of residential rooftop solar PV adoption have been widely researched and can be broadly categorized into four main categories: adopter characteristics (Mildenberger et al. 2019; Sunter et al. 2019), available economic incentives (Li and Yi 2014; Matisoff and Johnson 2017), regional spillover effects (Graziano et al. 2019; Graziano and Gillingham 2015), and urban form constraints

(Dharshing 2017). While a large body of work has developed strong theoretical foundations for determinants within the first three categories, less work has dedicated itself to exploring how urban form is characterized and interpreted in detail in such studies. In the following paragraphs, we describe several limitations in the current characterization of urban form in solar PV adoption analyses, the implications of this on the current interpretation of results, and an alternative means of representing urban form in such analyses.

4.2.1 Urban form characterizations in solar PV adoption analyses

Scholars focusing on urban form and rooftop solar adoption currently characterize urban form through housing density data or other variables highly correlated with housing density. The theoretical basis for quantifying urban form in this manner typically stems from two distinct areas: *solar rights* and *rooftop solar suitability*. *Solar rights* refers to the legal ability of a resident to install rooftop solar (Kettles 2008). Private or public covenants, ordinances, or building codes applicable to a residence may restrict or protect a resident's right to install solar PV, where restrictions tend to be placed on residence types that are either multifamily or not owned. Residents of multifamily housing types also incur barriers from misaligned benefits and costs between landlords and tenants, even when a resident can afford solar, commonly referred to as the split incentive problem (Gillingham et al. 2012). Because rented, multifamily properties tend to be concentrated in more urban, high-density areas, researchers have reason to expect higher density areas to have lower solar adoption rates, based on lower levels of solar rights. A number of studies have quantitatively found this to be the case in the contexts of Californian cities (Hsu 2018), Australian cities (Poruschi and Ambrey 2019), the United States (Kwan 2012), and Connecticut (Graziano and Gillingham 2015).

In contrast to solar rights, a few recent studies have reasoned that areas with lower housing density may also be more prone to solar adoption due to their increased *rooftop solar suitability* (Dharshing 2017; Hsu 2018; Poruschi and Ambrey 2019). Rooftop solar suitability refers to a building rooftop's suitability for solar based on its rooftop size, orientation, and shadowing from neighboring buildings or obstructions. In contrast to solar rights, which entails legal barriers to solar adoption, rooftop solar suitability is more relevant to the actual structure of the built environment. High rooftop solar suitability is indicative of urban forms designed in a way (whether intentional or not) that make solar PV systems more financially viable. In the context of this study, rooftop solar suitability is distinct from solar radiation; a neighborhood in Philadelphia may have the same rooftop solar suitability as a neighborhood in Florida, given a similar urban form. The combination of rooftop solar suitability, solar radiation, economic incentives, electricity price, and solar system cost compose the financial payback of the system. Rooftop solar suitability information is becoming increasingly accessible online and integrated into tools for consumers to make more informed decisions, such as Google Project Sunroof (Google 2020) and Sun Number ("Sun Number" 2020).

Rooftop solar suitability is in large part dependent on the roof's continued access to sunlight throughout the day. Conflicts and issues arising over solar access are not new, and policies for its protection existed long before concerns over solar panel return on investments. Originally sought for the purpose of protecting access to daylight, the right to sunlight was protected through the 'ancient lights' doctrine under English common law, which was generally adopted and practiced in the US until 1959 (Bronin 2009). In an effort to dissuade curbs in construction and development, a 1959 Florida court case rescinded the applicability of the ‘ancient lights’ doctrine, specifying landowners do not have a legal right to the light traversing across the land of their neighbors onto their own property (Pfeiffer 1982). Today, except in cases where private easement agreements are made, developments are generally not required to consider neighboring rooftop solar suitability in their design in the US; residents considering solar may not have enough rooftop solar suitability to begin with for the system to be financially viable or may be uncertain about their solar exposure in the future, both of which hinder solar adoption. Working under the assumption that residences in higher density areas have a higher likelihood of incurring shading from neighboring buildings or other obstructions, researchers have again suggested and quantitatively supported that higher density areas will have less solar adoption due to lower rooftop solar suitability (Hsu 2018; Poruschi and Ambrey 2019).

A summary of the aforementioned studies and their findings is provided in Table 1. While the results of such studies have amounted to relatively consistent results—lower density areas associated with higher solar adoption—it is less clear why this phenomenon is occurring. Research in this area has yet to distinguish between two foundationally divergent modes of reasoning, solar rights and rooftop solar suitability, as they are characterized by the same variable: housing density.

Table 11 – Summary of solar adoption research with focus on 'urban form'

Authors	Scope (Granularity)	Findings	Reasoning
Kwan, 2012 (Kwan 2012)	United States (Zip Codes)	- Housing density had a **negative** relationship with PV adoption.	- More housing units in higher density areas, creating less solar installations per housing unit
Graziano, 2015 (Graziano and Gillingham 2015)	Connecticut (Block groups)	- Housing density had a **negative** relationship with PV adoption. - Share of rented homes had a **negative** relationship with PV adoption.	- More split-incentive issues for rented housing, which is correlated with high-density areas
Hsu, 2018 (Hsu 2018)	California cities (cities)	- Housing density had a **negative** relationship with PV adoption.	- More solar approval processes and single-family housing in low-density areas - Less shading in low-density areas
Poruschi, 2019 (Poruschi and Ambrey 2019)	Australia cities (Postcodes)	- Share of apartments had **negative** relationship with PV adoption.	- More split-incentive issues for rented apartments - More daylight availability in low-density areas
Dharshing, 2017 (Dharshing 2017)	Germany (Counties)	- Single-family homes had **no significant** relationship with PV adoption.	- Relationship between single-family homes and solar adoption remains unclear
Graziano, 2019 (Graziano et al. 2019)	Connecticut urban case study (Block groups)	- Housing density had **no significant** relationship with PV adoption. - Share of rented homes had **no significant** relationship with PV adoption.	- Relationship between housing density and solar adoption and rented homes and solar adoption remains unclear

Understanding whether high density areas tend to have limited solar adoption because of predominantly solar rights constraints or a combination of solar rights and rooftop solar suitability issues, the degree to which each of these factors influences adoption, and how these factors stand in contrast to other known determinants of residential solar adoption have important implications for solar energy policy and research. If lower solar adoption in high density (i.e., urban) areas is being primarily driven by legal restrictions in multifamily or rented units, the set of interventions necessary to address these barriers are unique and distinct from the interventions needed to address rooftop solar suitability constraints. Previous research has investigated in detail solutions that promote solar rights and incentives that address barriers for solar adoption in multi-family contexts (Sustainable Energy Roadmap 2016). Meanwhile, the policies and solutions that facilitate rooftop solar suitability in high density areas are less clear, likely in part due to our lack of understanding regarding the degree to which rooftop solar suitability, or urban form, influences solar adoption. For example, it is currently unclear how influential neighboring shading is (Xu et al. 2014) and how it may affect solar decision making relative to other factors known to influence solar uptake. If a strong relationship exists between rooftop solar suitability and solar adoption, this would indicate that the physical structure of a city also influences solar decision making and uptake. Identifying this would broaden our understanding of solar adoption trends and add more nuance to the previously found housing density results by disentangling its confounding effects with solar rights. Moreover, knowledge regarding the importance of rooftop solar suitability across our built infrastructure can provide important insight for policy strategies and urban design guidelines that aim to support new populations and renewable energy integration across our cities. However, a major challenge in the literature in discerning the relationship between rooftop solar suitability and solar adoption is defining urban form in a way that is distinct from housing density.

4.2.2 Defining urban form through rooftop solar suitability

Rooftop solar suitability has been studied and defined using more nuanced constructs (i.e., beyond housing density) in architecture and engineering domains. Of note, these domains have used different terminology to describe similar concepts to rooftop solar suitability, using terms such as *solar potential* or *solar access*. For clarity and consistency, this paper will continue to refer to these same concepts using the term rooftop solar suitability throughout. Scholars have examined how rooftop solar suitability across built infrastructure varies in relation to different urban forms, identifying influential factors including: building density, building height, building orientation, site coverage, street layout, and rooftop areas, as well as other planning conditions (Lobaccaro et al. 2017; Sanaieian et al. 2014; Sarralde et al. 2015). While these research efforts started out using theoretical urban simulations to optimize rooftop solar suitability (Cheng et al. 2006; Hachem et al. 2011), the increase in volume and accessibility of city data (i.e., 3D models, satellite imagery, LiDAR data) and computational power has led to emerging research efforts to accurately quantify existing and predict future rooftop solar suitability across cities around the world (Lobaccaro et al. 2019; Sigrin and Mooney 2018; Zhu et al. 2019). In one analysis, researchers quantified rooftop solar suitability across different urban block typologies; under the same planning conditions, an increase in 200% rooftop solar suitability could be achieved based on different urban block designs (Zhang et al. 2019). Likewise, another study found that rooftop solar suitability could increase by 25% for the case study of buildings examined (Lobaccaro et al. 2017). Such studies aim to inform architects and urban planners of best practices and driving factors for expanding rooftop solar suitability in the construction of new neighborhoods or development of existing neighborhoods. They also work under the assumption that greater rooftop solar suitability will lead to greater solar adoption, although this relationship has yet to be empirically investigated.

The broader research question in this study asks: does urban form impact solar PV adoption? We explore this research question by assessing one metric representing urban form, rooftop solar suitability, and examining the relationship between rooftop solar suitability and solar PV adoption levels. As areas with more rooftop solar suitability can produce more energy from solar PV, improving the financial payback of the system, we propose the following hypothesis for this study:

- **Hypothesis**: Areas with more rooftop solar suitability will have higher solar PV system adoption rates.

Our analysis seeks to build upon previously published literature in several ways. Previous work has relied on measurements of housing density to quantify urban form's impact on solar adoption (Dharshing 2017; Hsu 2018; Poruschi and Ambrey 2019). We use measurements of rooftop solar suitability to more directly represent urban form and to disentangle the effects of rooftop solar suitability from solar rights. We also quantify the influence of rooftop solar suitability in relation to other known factors affecting solar adoption (e.g., political leanings, financial incentives, socio-demographics). This contextualization informs the importance of rooftop solar suitability relative to other factors already documented in the literature. Finally, we connect the conceptualization of rooftop solar suitability, as defined in architecture and engineering fields, to solar adoption analyses found in policy domains.

4.3 Methods

This study considers residential solar PV system counts in the analysis. The geographic scope for this analysis is a previously understudied area in terms of solar adoption analyses—Metropolitan areas across nine states comprising the Southeastern US: Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, and Tennessee (Figure 1). The cutoff for areas defined as metropolitan follows thresholds set by the US Census (US Census 2019). Regression analysis techniques are used to evaluate the hypothesis. In the sections below, the variable of interest, dependent variable, and control variables are described first. Following this, the regression modeling and validation approach are detailed.

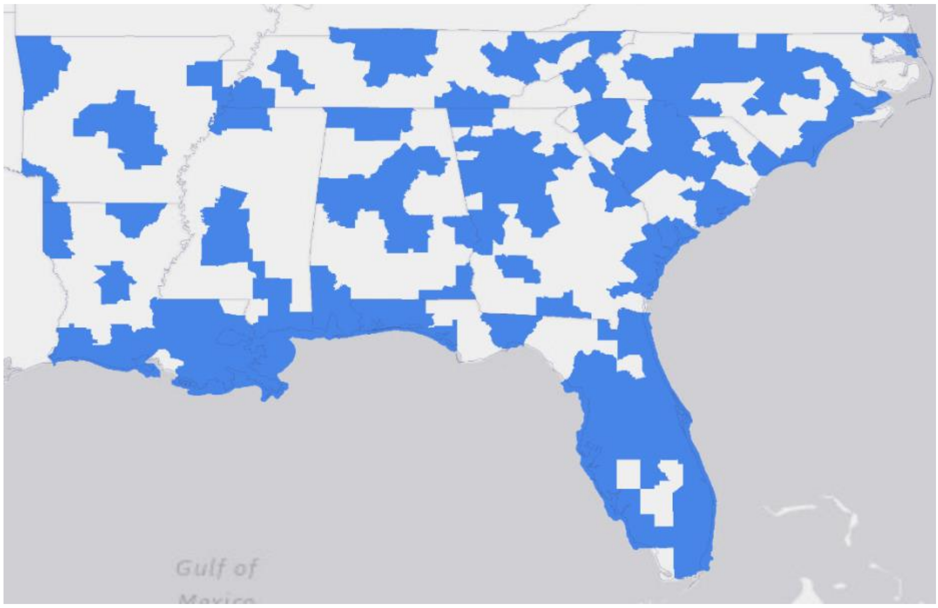


Figure 12 – Study geographic scope includes metropolitan areas in the Southeastern US (represented by the blue shaded regions)

4.3.1 Data

4.3.1.1 Rooftop solar suitability Data

The primary variable of interest is the *rooftop solar suitability* of each census tract. Rooftop solar suitability represents the suitability of rooftops for solar energy generation (in this study considering only residential rooftops), accounting for infrastructure constraints such as rooftop area, tilt, or neighboring shading. In this study, the data for rooftop solar suitability was developed by Gagnon et al. and Sigrin and Mooney (Gagnon et al. 2016; Sigrin and Mooney 2018), who computed rooftop solar suitability of each census tract, i, by the equation:

$$Rooftop\ Solar\ Suitability_i = \frac{Count\ of\ residences\ suitable\ for\ solar_i}{Count\ of\ total\ residences_i} \tag{4}$$

The numerator, the count of residences suitable for solar, is a function of four requirements: roof plane shading, orientation, tilt, and area (Table 2). If any of the four requirements were not met, a building was classified as not suitable for solar. For example, a south-facing rooftop that was 20 m^2 in area, had a 45 degree tilt, but was located in a city where it only had 40% illumination during December would be classified as 'not suitable'. Sigrin and Mooney (Gagnon et al. 2016; Sigrin and Mooney 2018) used LIDAR data from 128 US cities combined with PV-generation modeling to quantify each requirement from Table 2. Both single and multifamily buildings are included in the residence count in Equation 4.

Table 12 – Mandatory requirements for rooftop to be considered 'suitable' for solar (Adapted from Sigrin and Mooney 2018)

Category	Requirement(s)
Shading	60% illumination (March, September) 50% illumination (December) 70% illumination (June)
Orientation	Rooftop surface faces either: E, SE, S, SW, or W
Tilt	Average surface tilt is less than 60 degrees
Minimum Area	At least 10 square-meters

Importantly, rooftop solar suitability was normalized given a residence's geographic location, and it is *not* dependent on the level of solar radiation it receives. In other words, rooftop solar suitability in this study only accounts for the physical infrastructure of the building itself and the effects of neighboring buildings; it does not vary based on a building being located in an area with more or less solar radiation (e.g., North Carolina vs Florida). This data provides a more nuanced characterization of rooftop solar suitability beyond housing density; namely, it contends with the morphology of a building's surrounding environment and its individual roofing characteristics. A summary of the distribution of the rooftop solar suitability across residential rooftops is provided in Figure 2 (distributions by state can be referred to in the Supplementary Files S1).

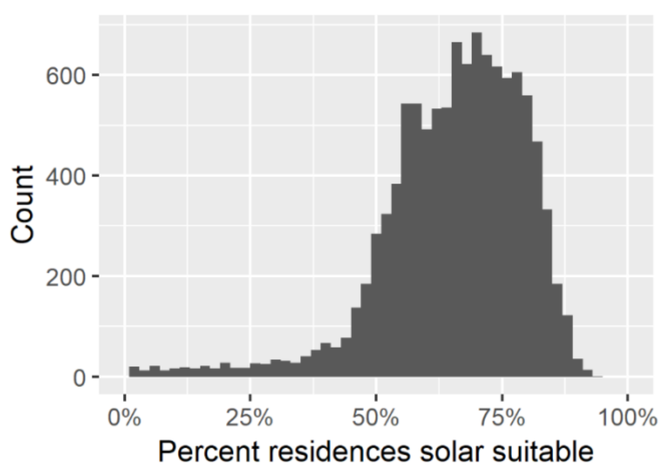


Figure 13 – Distribution of percent residences suitable for solar by census tract

4.3.1.2 Solar PV system count data

The dependent variable for the analysis is the count of residential solar PV systems per census tract. This data was obtained from the DeepSolar database, a recently released public database containing solar PV array locations throughout the US as of December

2017, aggregated by census tract (Yu et al. 2018). The solar PV system counts in DeepSolar were detected from satellite imagery using a convolutional neural network deep learning framework. This dataset stands in contrast to traditional means of obtaining solar PV data; most previous studies have depended on either the US-based Open PV Project or Tracking the Sun (Hsu 2018; Kwan 2012; Li and Yi 2014; Lukanov and Krieger 2019), or state or country-wide utility programs that collect data during the duration of a particular program (Dharshing 2017; Graziano et al. 2019; Poruschi and Ambrey 2019). The Open PV Project and Tracking the Sun have thorough documentation of solar systems installations on the West Coast and Northeastern US; however, they lack representation in the Southeastern US. Currently, such databases have no recorded solar installations in the Southeastern US with the exception of Florida. The DeepSolar database overcomes this geographical bias by not relying on reporting from state agencies, and while its accuracy may be inhibited by other means (e.g., misidentified or overlooked panels), solar system count representation is maintained in regions that have been largely overlooked by solar research thus far. A summary of the distribution of solar PV systems per census tract is provided in Figure 3 (distributions disaggregated by state can be found in Supplementary Files S2).

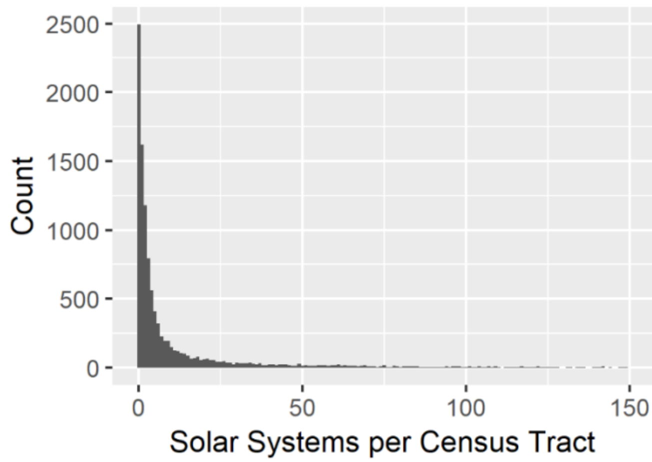


Figure 14 – Distribution of solar PV systems per census tract

4.3.1.3 Control variables

Many demographic, environmental, and policy factors have been consistently identified to be associated with solar PV system adoption. Table 3 lists the factors used as controls in this analysis, which are in line with variables frequently applied in other analyses. The raw data downloaded from respective sources listed in the table was used in the analysis, with the exception of the Solar Policy variables. Similar to Crago and Chernyakhovskiy (Lasco Crago and Chernyakhovskiy 2017), we accounted for variations in solar policies by controlling for the binary presence of solar incentives offered in each state. State-level solar policies were recorded from the DSIRE database, and disaggregated into six types of policies: cash incentives, property tax incentives, sales tax incentives, tax credits, Renewable Portfolio Standards (RPS), and net-metering policies (Matisoff and Johnson 2017). If a policy had been enacted during 2017 or prior, the census tracts within that state received a one associated with the specific policy type. States that did not enact the type of policy as of 2017 received a zero. Policy types were included in the analysis if they had been enacted in at least two states across the sample period to provide adequate variation in the policy predictors.

Table 13 – Control variables used in analysis

Variable	Quantification	Source
Population (log)	Number people	ACS (ACS 2016)
Housing Density (log)	Average number of housing units per square-meter of land area	
Income (log)	Average household income	
Rented housing	% housing units rented	
Education (sqrt)	% population with a bachelor's degree	
Political Affiliation	% population Democrat	
Employment	% population employed	
Race- White	% population White	
Solar Radiation	Daily average Watt-hours per m^2	NASA (NASA 2008)
Electricity Rate	Average residential retail rate ($/kWh)	NREL (NREL 2017)
Solar Policies	Property tax incentive (binary)	DSIRE (DSIRE 2019)
	Tax credit (binary)	
	Net-metering (binary)	

4.3.1.4 Data collection and transformation process

The DeepSolar database contains the data from the ACS and NASA sources specified in Table 3. We merged the DeepSolar database with data from the NREL (NREL 2017; Sigrin and Mooney 2018) and DSIRE (DSIRE 2019) databases based on the 11-digit FIPS code. Next, we filtered the data to include only census tracts located within Metropolitan Statistical Areas in the Southeastern US, as defined by US Census thresholds (US Census 2019). The remaining number of census tracts totaled 11,917. Observations were dropped when null values were present in any of the variables. A total of 973 observations were dropped, 826 of which were due to a null value for the solar radiation data. A total of 10,944 observations remained. All data cleaning and filtering was performed using Python 3.7.

4.3.2 Statistical models

4.3.2.1 Hypothesis testing

To investigate the relationship between rooftop solar suitability and solar adoption, the following empirical model was designed:

$$\boldsymbol{log(PVcount_i) = \beta_0 + \beta_1 RooftopSolarSuitability_i + \beta_2 X_i + \varepsilon_i}$$

where $PVcount_i$ is the number of solar PV systems installed in census tract *i*; $RooftopSolarSuitability_i$ is the percent of residences suitable for solar in census tract *i*; X_i is a vector of control variables from Table 3 likely to be associated with solar adoption; and ε_i is the error term.

The response variable, solar PV system count, is a non-negative integer with a distribution that is clearly non-Gaussian (Figure 3). Applying OLS regression methods in this case is inappropriate as the high number of zeros prevent the transformation of the response variable to a more normal distribution, and any negative predictions using this approach would have no theoretical basis (i.e., there cannot be negative solar systems counts) (O'Hara and Kotze 2010). Count regression models, including Poisson and its variants, were considered next. Poisson assumes the variance of the response is approximately equal to its mean. In this case, the mean number of solar PV systems counted across all census tracts was 15.6, with a variance of 1857. Thus, the assumptions for Poisson regression are not met, and furthermore, over-dispersion is highly evident (Hilbe 2011). In contrast to Poisson regression, which uses a single term to represent the mean and variance, negative binomial regression incorporates a separate parameter for each allowing for more flexibility in modeling when the outcome variable is over-dispersed. For this reason, we used negative binomial regression to model solar system counts as a function of its predictors, which has been applied across several solar adoption analyses (Hsu 2018; Kwan 2012; Lasco Crago and Chernyakhovskiy 2017)[4].

Multi-collinearity was examined between explanatory variables using correlation matrices and by computing the Variance Inflation Factor (VIF) for each variable within the model. Coefficient p-values below 0.05 were designated as indicating statistical significance. Models denoted as containing standardized covariates indicate that all continuous covariates were scaled to have a mean of zero and standard deviation of one. All statistical analyses were conducted using R Programming language version 3.5.1.

4.3.2.2 Validation approach

In the absence of panel data, we employed several tests to assess the robustness of the rooftop solar suitability coefficient. First, we observed the coefficient movements as we gradually added controls to the negative binomial model. The stability of the coefficient across models with varying controls can be an indicator of the coefficient's sensitivity to omitted variable bias. In this assessment, we examined movements in the pseudo R-squared

[4] Zero-inflated negative binomial models were also considered; however, they were not used in the analysis because of the lack of a strong theoretical basis that two separate processes are generating zero counts.

values as we added controls to consider the importance of the controls relative to the coefficient stability (Oster 2019). Second, the policy variables were transformed to explore rooftop solar suitability sensitivity to alternative quantifications of policy variables. Instead of including a binary term for each policy, the policies were aggregated to one variable, indicating the total number of state-wide solar policies implemented in each state, similar to Kwan (Kwan 2012). Third, quantile regression was implemented to evaluate the rooftop solar suitability coefficient sensitivity to varying levels of the response variable. By quantifying covariate coefficients against different percentiles of the response variable, quantile regression enables a more complete understanding of the association of a covariate across the entire distribution of the response variable (Koenker and Hallock 2001). We used the *quantreg* package in R (Koenker 2019) and modeled covariate coefficients by setting tau (the sample quantile) in increments of 0.10. We specifically examined movements in the rooftop solar suitability coefficient across quantiles.

4.4 Results

4.4.1 Hypothesis test results: relationship between rooftop solar suitability and solar adoption

Table 4 contains a list of covariate coefficients and their significance. Standardized covariates are provided in the second column to ease comparisons between coefficient magnitudes. Rooftop solar suitability was found to be positively and significantly (p-value <0.001) associated with solar adoption, enabling the rejection of the null hypothesis. The interpretation of negative binomial model results is as follows: the rooftop solar suitability coefficient implies that for a one-unit change in rooftop solar suitability, the difference in the log of the expected count of solar PV systems is expected to increase by 4.627. Comparing the standardized coefficients, rooftop solar suitability's association with solar adoption (0.666) is slightly lower than that of solar radiation (0.704), and higher than solar adoption's association with the remaining variables, including: population (0.381), housing density (0.227), education (0.206), income (0.201), political affiliation (0.164), race (-0.144), rented housing (0.114), employment (0.065), and electricity rates (0.055).

Table 14 – Negative Binomial Regression Results

	Raw covariates	Standardized covariates
Log(population)	0.72*** (0.025)	0.381*** (0.013)
Solar radiation	2.221*** (0.064)	0.704*** (0.02)
Solar suitability	**4.627*** (0.153)**	**0.666*** (0.022)**
Log(income)	0.468*** (0.068)	0.201*** (0.029)
Sqrt(education)	1.263*** (0.155)	0.206*** (0.025)
Employment	1.105*** (0.301)	0.065*** (0.018)
Political affiliation	1.037*** (0.134)	0.164*** (0.021)
Race- White	-0.537*** (0.074)	-0.144*** (0.02)
Log(housing density)	0.142*** (0.012)	0.227*** (0.019)
Rented housing	1.009*** (0.245)	0.114*** (0.028)
Electricity rate	4.828*** (1.337)	0.055*** (0.015)
Net metering policy[a]	0.668*** (0.054)	0.668*** (0.054)
Property tax policy[a]	1.197*** (0.035)	1.197*** (0.035)
Tax credit policy[a]	-0.076 (0.042)	-0.076 (0.042)

*Statistical significance: ***(p < 0.001), **(p < 0.01), *(p < 0.05); [a] Variable is binary*

To check for multicollinearity, the maximum variance inflation factor across all the variables in the model was 5.3. The average variance inflation factor for the model was 2.67. In a post-hoc analysis, the regression coefficients were computed using the rate of solar system adoption, as opposed to the solar system count, as the dependent variable. The rate of solar system adoption was defined as solar system count per residence count in each census tract. To model this with a negative binomial approach, a residence count offset term was added to the model and the population term was removed from the model. The results can be referred to in the Supplementary Files S3. The coefficient directions and significance remained consistent compared to the solar system count model in Table 4.

4.4.2 Validation results

The validation approach assessed the stability of the rooftop solar suitability coefficient across the choice of covariates, policy representation, and solar adoption quantiles. Table S4 in the Supplementary Files contains the rooftop solar suitability coefficient when adding groups of covariates. Across all models (columns 1-5), rooftop solar suitability remains positively and significantly (p-value <0.001) associated with solar adoption. Its coefficient magnitude is consistently equal to or greater than the majority of controls across the model. The model containing the alternative representation of policy variables (column 5) also gives consistent coefficient results regarding rooftop solar suitability.

The results of the quantile regression are presented in Figure 4. Each point represents the value of the coefficient at a certain quantile, with the vertical bar representing the standard error for the coefficient. No coefficient exists for the 0.1 and 0.2 quantile as all dependent values are zero. Rooftop solar suitability remains statistically significant and positive for all remaining quantiles. The standard error tends to increase at higher quantiles, which is expected given the variance of the outcome variable increases at higher quantiles.

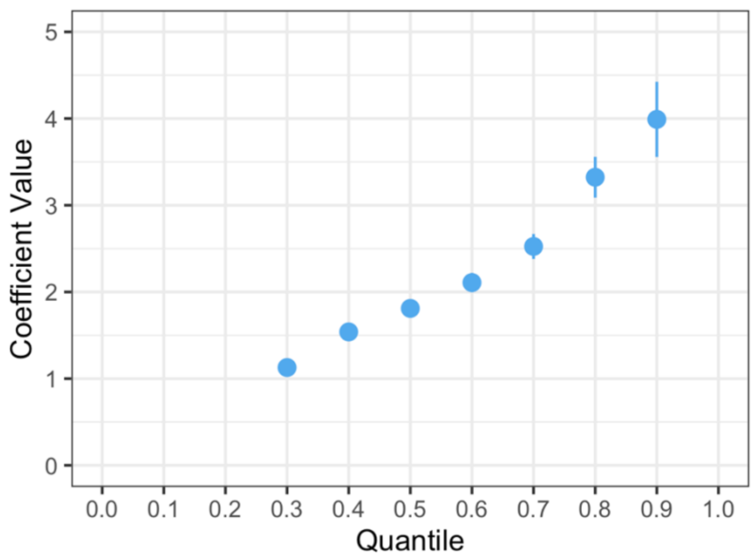


Figure 15 – Rooftop solar suitability coefficient values across varying quantiles of solar adoption

4.5 Discussion

The results of this study show a positive and significant association between rooftop solar suitability and solar adoption. This provides evidence to reject the null hypothesis, and indicates that census tracts with more residential rooftop solar suitability tend to have greater residential solar adoption in the Southeastern US. While these findings may be intuitive, solar adoption analyses have largely lacked incorporation of rooftop solar suitability metrics, or other direct measures of urban form. Instead, considerations of urban form have relied almost entirely on housing density metrics (Hsu 2018; Poruschi and Ambrey 2019). Solar adoption analyses have theorized urban form structure to be an important influencer of solar adoption (Dharshing 2017), yet discerning meaning from housing density metrics is convoluted; housing density contains clear confounding factors unrelated to urban form. The most prominent of these is the impact of solar rights (i.e., areas with less housing density have more single-family, owned homes). Thus far, research has yet to examine how the built infrastructure of a city may be correlated with solar adoption decisions in a way that is disentangled from solar rights. The results of this study contribute to solar adoption literature by using a more direct measure for the shape of built infrastructure, rooftop solar suitability, and connecting this with solar adoption while controlling for other explanatory factors.

Part of the reason more direct metrics for urban form, such as rooftop solar suitability, have not been integrated into solar adoption analyses is the lack of comprehensive and consistently measured data on the built environment. Scholars in engineering and architecture domains have attempted to quantify rooftop solar suitability for decades (Cheng et al. 2006; Hachem et al. 2011) with data based on theoretical models, which lacked empirical metrics for real cities. In more recent years, this data has become more accessible across large geographic scales with high granularity in real environments (Lobaccaro et al. 2019; Sigrin and Mooney 2018; Zhu et al. 2019). Yet, incorporation of rooftop solar suitability data into solar adoption analyses is far from common practice. This is especially poignant as our findings show rooftop solar suitability has a similar magnitude of association with solar adoption (0.556) compared to other fundamental inputs into solar modeling, such as solar radiation (0.640). Furthermore, the application of these new data sources (Sigrin and Mooney 2018; Yu et al. 2018) enabled a more up-to-date analysis through 2017[5] in a context where empirical studies are nearly non-existent, the Southeastern US.

[5] Of the studies providing the basis of the literature review in this study, four studies had data through 2013 (Dharshing 2017; Graziano et al. 2019; Graziano and Gillingham 2015; Hsu 2018), one had data through 2010 (Kwan 2012), and one had data through 2016 (Poruschi and Ambrey 2019).

While these results cannot be interpreted as a causal link, it is encouraging that the covariate results were largely coherent with previous assessments. For example, increases in a census tract's population (Hsu 2018), solar radiation (Dharshing 2017), income (Poruschi and Ambrey 2019), education (Hsu 2018; Kwan 2012), democratic political affiliation (Kwan 2012), and electricity price (Hsu 2018; Kwan 2012) have all been found to be significantly associated with an increase in solar adoption, which also aligns with intuitive expectations. Housing density was significantly and positively correlated with adoption, when controlling for rooftop solar suitability and rented housing. This stands in contrast to the many studies that find housing density, or its proxies, to be negatively correlated with adoption (Graziano and Gillingham 2015; Hsu 2018; Kwan 2012; Poruschi and Ambrey 2019) or have no significant correlation (Dharshing 2017; Graziano et al. 2019). It has been theorized that housing density is correlated with multiple competing factors—namely rooftop suitability for solar, solar rights, and peer effects (Dharshing 2017). With rooftop suitability and solar rights controlled for, we suggest that the positive coefficient is capturing mainly peer effects, where solar technologies are theorized to be more likely to be adopted by consumers with proximal neighbors who adopt solar (Dharshing 2017; Graziano and Gillingham 2015). In addition, local solar initiatives such as Solarize programs were not accounted for in this analysis, but tend to be located in cities—where housing densities are higher. The positive relationship between housing density and adoption may also be a reflection of local solar initiatives concentrated higher density areas.

Some variables presented more nuanced findings. The percentage of people identifying as white in a census tract had a significant and negative relationship with solar adoption. Of the few studies that have considered racial identities in analyzing solar adoption trends, most have found lower levels of solar adoption in areas with higher levels of people identifying as belonging to an ethnic or racial minority group (Bollinger and Gillingham 2012; Graziano et al. 2019; Sunter et al. 2019). These studies have been conducted in the context of Connecticut (Graziano et al. 2019), California (Bollinger and Gillingham 2012), and across the US (Sunter et al. 2019). The alternative trend found in this study may be a result of the new geographical context under consideration. To examine this phenomenon in greater detail, household-level data is necessary. Additionally, rented housing was found to be positively and significantly associated with solar adoption. This stands in contrast to studies that have found a negative relationship between apartments or rented housing and solar adoption (Poruschi and Ambrey 2019). These findings have been based on the premise that installing solar is often more challenging for those living in a multifamily complex and makes less economic sense as a renter. The magnitude of the association found in this study is relatively low compared to other covariates. As with all analyses with aggregated data, there is potential for different trends to exist at the individual level compared to trends at the group level (i.e., the ecological fallacy). In this case, the census tracts with higher levels of adoption associated with higher levels of rented housing may in fact be due to single family homes within the census tract adopting solar. Nonetheless, it is an intriguing finding given that housing density and rooftop solar suitability are also controlled for in the model.

The policy results show a positive and significant relationship between net metering and property tax policies and solar adoption, and a negative and non-significant relationship between tax credit policies and solar adoption. Our policy metric shortcomings and potential implications are discussed in further detail in the limitations section. In short, the impact of policies is coarsely represented, and we express caution in the reliable interpretation of their coefficients in terms of shedding light on the actual policy effectiveness. The purpose of these metrics is to examine rooftop solar suitability's sensitivity to representations of policy variables. Across models without policy variables, with individual policy variables, and with aggregated policy variables (Supplementary Files S4), the results for rooftop solar suitability remained positive and highly significant. This gives us confidence in the stability of the results and opens avenues to future research to move forward with studying this relationship in more detail.

4.6 Implications

4.6.1 Policy implications

Our findings, and future efforts examining urban form and solar adoption, have important implications for cities promoting solar adoption and renewable energy uptake. The results suggest there is the potential to increase and promote solar adoption, or conversely the risk of retracting it, as our infrastructure is adapted, infilled, and evolved to meet future needs. Greater attention to the dependencies between infrastructure systems, for example how densification may compromise rooftop solar suitability, is needed by designers, planners, and engineers. Such disciplines should have the technical background to assess the rooftop solar suitability across a network of buildings, develop innovative design solutions for maintaining rooftop solar suitability, and critically examine trade-offs in achieving rooftop solar suitability versus other sustainability pursuits. Coordinated approaches to rooftop solar suitability will also require involvement and support from local or state governments. Rooftop solar suitability is composed of multiple design factors (i.e., shading, tilt, orientation, and rooftop area), which policies could more directly address. There are a number of pathways that have been used for promoting rooftop solar suitability at the local level, primarily: zoning rules, access permits, and easements (SF Environment 2012), with varying degrees of protections for solar adopters or property restrictions for solar neighbors. A natural avenue for future research is to identify the policies that are most effective at increasing rooftop solar suitability. In adopting rooftop solar suitability policies and integrating rooftop solar suitability into design guidelines, we are careful to not imply that a single urban form structure exists that will optimize rooftop solar suitability. Strategies will be highly dependent on the existing urban morphologies, as well as the norms and desires of a community.

4.6.2 Broader research implications

More broadly, this study aims to connect disciplines that have complementary pursuits but are traditionally siloed, which previous research has called for (Sovacool 2014); in this case, we explicitly draw new links between computational analytics in engineering and architecture disciplines and solar adoption methodologies in public policy and social science domains. This process has expanded insights at the intersection of solar adoption analyses and built infrastructure systems. While solar adoption analyses tend to focus on adopter characteristics and available economic incentives (Schelly 2014), city infrastructural constraints—or opportunities—and their potential influence on solar decisions is less understood. Infrastructure, as determined by built environment policy, is long-lasting. Studying interactions and unforeseen outcomes between the shape of the built environment and available climate change solutions has been brought to attention by scholars in recent years as a pressing research area that is under-addressed (Creutzig et al. 2016). In the context of city infrastructure and solar PV, infrastructure densification has become a prominent strategy for cities as they evolve to accommodate new populations and foster sustainable urban planning (e.g., transit-oriented development). Densification resulting from urban growth in cities will likely increase shading from neighboring buildings, blocking the financial viability for many potential solar PV systems. Compounding this barrier to adoption is the risk consumers take on when installing a solar PV system that may be shaded by future developments, as the right to sunlight is not protected in most jurisdictions. Prior to this study, the importance of urban form on solar decision making lacked in-depth exploration. This study takes a step to build our understanding of the connection between urban form design and one major climate change solutions area—solar PV adoption. While we express caution in interpreting these results as definitive, the positive and significant correlation found between rooftop solar suitability and solar adoption is a new and notable relationship, and provides a foundation for future work to explore this phenomenon further.

4.7 Limitations and future research

The primary limitation of this study is that the solar panel adoption and rooftop solar suitability data do not contain information for each census tract across different time periods. Endogeneity bias from omitted explanatory variables that are correlated with the covariates is likely to be present due to the inability to explore intra-census tract variations. The study does evaluate the stability of the coefficient of interest by intentionally removing control variables and evaluating its value at varying quantiles of the response variable. In both scenarios, it is reassuring that the results for the rooftop solar suitability coefficient remained consistent. We do not seek to claim that results demonstrate a causal relationship between rooftop solar suitability and solar adoption; rather, the study seeks to draw connections between two related fields and establish new evidence for future research to build upon with additional robustness checks. One area for future work is to collect time varying solar PV adoption data by implementing open source algorithms such as the DeepSolar deep learning algorithm (Yu et al. 2018) on historic and current satellite imagery. Google Project Sunroof's (Google 2020) database may also be a fruitful data source for both solar PV adoption and rooftop solar suitability data, depending on if it is periodically updated with new data. As of the writing of this paper, Google Project Sunroof does not release this data at multiple time periods.

The analysis could also benefit from incorporation of policy information that quantifies the magnitude of each policy, as well as local or sub-statewide policies. While inclusion of these more detailed metrics will likely change the value of rooftop solar suitability's coefficient, we do not expect the sensitivity of the coefficient to be sufficiently extreme to change the significant findings of this coefficient. The addition of policy variables in the analysis did not substantially change the magnitude of the coefficient of interest, or its significance. Future studies may also consider conducting a similar analysis across a smaller geographic region to capture deviations in local level policy incentives across census tracts. The Southeastern US has a specific context for solar adoption; conducting this analysis in different geographic, infrastructural, and political contexts may yield interesting and alternative insights. There is also potential to add interactions terms representing interactions between different policy types to represent policy mechanisms with greater nuance. We tested models including interactions between policy variables, however these interactions did not meaningfully change the results for the rooftop solar suitability coefficient.

Finally, given urban form likely has an impact on solar adoption, a natural extension of this analysis is examining the policies that shape urban form to be more or less suitable for solar. A region's rooftop solar suitability ultimately depends upon the policies, zoning, and building codes that dictate the design of its infrastructure and access to sunlight. It would be informative for future studies to examine drivers of rooftop solar suitability from a policy perspective. A potentially interesting consideration in this research direction would be consideration of soft costs, which are a known barrier to solar adoption (Hsu 2018). Some rooftop solar suitability policies may be time-consuming, such as negotiating assurances to access sunlight through an easement, which in effect could limit the ability of residents to take advantage of a policy even if it is enacted.

4.8 Conclusion

Urban forms, as determined through built environment policy, are long-lasting. As cities adapt their infrastructure systems to become more livable, sustainable, and resilient, it will be important to understand the secondary impacts of such changes. This paper examines the relationship between different urban forms and solar PV adoption to identify if and how changing urban forms can encourage or detract from the viability of solar PV systems in the future for cities. Solar adoption analyses have predominantly concentrated on how adopter characteristics, available incentives, and regional spillover effects are associated with adoption. While such analyses have considered the impact of urban form, they have lacked a deeper reflection on how urban form was characterized, leading to unclear implications of their findings. Inspired by recent progress in engineering and architectural disciplines to develop rooftop solar suitability information, we apply this new data to more directly assess the relationship between urban form and solar PV adoption in metropolitan areas in the Southeastern US. The results show a positive and significant relationship between rooftop solar suitability and solar PV adoption. This insight draws previously unexplored connections between solar adoption analyses and engineering computational work, and provides a basis upon which future studies can build at the intersection of these two areas. Improving our cities will require greater understanding of the dependencies between our infrastructure systems and trade-offs between best practices. The results of this study provide a crucial empirical foundation in furthering this understanding in the pursuit of sustainable and resilient cities.

4.9 Acknowledgements

This material is based upon work supported by the National Science Foundation under Grants No. DGE-1650044 and CPS-1837021. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

CHAPTER 5. CONTRIBUTIONS

Across my doctoral research, I led efforts at the intersection of energy data analytics, human-computer interaction, and public policy analysis to transform new forms of community energy data to be more understandable, engaging, and actionable to specific stakeholders. As data-driven, community-scale sustainability initiatives continue to attract attention, evaluating methods seeking to improve the usability of the myriad of data sources will become increasingly important (Zhou et al. 2016).

To this end, I began my doctoral research with an examination of the use of electricity information extracted from advanced metering infrastructure (AMI) installed across the Georgia Tech campus' buildings. Using building energy analytics methods, I developed new metrics geared towards informing retrofit decision making for building portfolio owners and municipalities. Expanding on the usability of this data analytics approach, my next study concentrated on the visualization of this same campus-scale energy data. I developed a mobile application that transformed building energy consumption and production data into a visual interface and studied its potential to engage Georgia Tech community members. The final study took a broader approach and studied energy decision making at the urban-scale. The analysis drew from literature in both public policy and computational engineering/architecture domains to examine relationships between solar PV uptake and built infrastructure form. The results of this analysis are geared towards urban planners to better inform how the design of the built environment may impact the viability of clean energy resources. In the following sections, the contributions of each study and opportunities for future research based on this study are discussed.

5.1 Temporally segmented building energy benchmarks

I began my research by reviewing current trends in building energy benchmarking and smart meter analytics. While building energy benchmarking studies in recent years have focused on applying new statistical and machine learning techniques to improve annual benchmark accuracies, little work in this area has focused on integrating more detailed energy data, such as smart meter information (Buck and Young 2007; Kavousian and Rajagopal 2014; Zhang et al. 2015). Smart meters are gaining prevalence across the commercial building sector (EIA 2018); entities from building portfolio owners to municipalities with access to this information have potential to make more responsive and contextual energy management decisions if this data can be transformed into actionable metrics. In this analysis I leveraged smart meter electricity data for buildings across Georgia Tech's campus and applied regression-based benchmarking techniques to develop a metric I refer to as temporally segmented building energy benchmarks. These new metrics have several academic and practical contributions.

While metrics comparable to temporally segmented building energy benchmarks have been calculated in previous works (ElYamany et al. 2017; Roth and Jain 2018), they were not the focus of these analyses, and had not been statistically compared to a control group. The hypotheses conducted in this analysis involved comparing distributions of temporally segmented and non-temporally segmented benchmarks (i.e., conventional energy benchmarks), and the results showed statistically significant differences between the two for the majority of buildings and temporal periods studied. These results contribute to the building energy benchmarking literature by implementing a robust statistical evaluation that quantifies differences between temporally segmented benchmarks and a control group. These differences are indicative that temporally segmented building energy benchmarks can enlighten additional information about building operations that was previously unknown.

This study also draws new connections between building energy benchmarking (ElYamany et al. 2017; Francisco et al. 2018a; Roth and Jain 2018) and smart meter analytics fields (Wang et al. 2018)—both domains of which are working towards improving building energy management decisions to increase the energy efficiency of the built environment. The results demonstrate that temporally segmented building energy benchmarks can show previously masked periods of inefficiencies. Practically, the results can be used for both: (1) identification and prioritization of specific retrofit strategies, and (2) near-time building energy management. For example, addressing building inefficiencies during time periods such as summer-peak periods and non-operating hours likely requires different retrofit approaches. Differentiating between building energy efficiencies during strategic time periods such as these can help building portfolio managers and municipalities prioritize which building(s) should receive an air-conditioner upgrade first, for example, to achieve the greatest energy savings. In addition, computing building energy benchmarks on a more granular basis can help enable quick identification of deviations in a building's performance, in the context of how the rest of the buildings in a community are performing. Of interest to building owners and cities is reducing energy use, optimizing allocation of resources, and making responsive and contextual decisions.

Temporally segmented building energy benchmarks are one option to help strive towards achieving these goals.

Looking ahead to future research in this area, a natural extension of this research effort is further validating temporally segmented building energy benchmarks by observing changes before and after a building retrofit. Ample research questions could be asked through these observations such as: Does building energy efficiency increase after an energy retrofit is implemented? Does improvement in energy efficiency occur during particular time periods? How does the type of retrofit affect these results? Are there implications for demand-side management based on the time periods when efficiencies are realized? This endeavor is being partly explored by another researcher in the lab; however, a major challenge in developing answers to these questions is collecting retrofit information consistently and having a large enough sample size of buildings with similar retrofits. In addition, the robustness of the developed benchmarks could be examined by applying different benchmarking methodologies, such as those with machine learning approaches. Similar strategies have been used to assess the validity of conventional building energy benchmarks (Li et al. 2014).

5.2 Community-scale energy feedback systems

Following the energy benchmarking approach to further the utility of campus energy information, the next step in my research involved assessing approaches for visualizing and engaging the public with the same energy data. City and community policies are increasingly requiring the public reporting of commercial building energy information (Institute for Market Transformation 2019). This data has potential to make the energy performance of buildings across communities more transparent, when historically information on buildings' energy impact—and embedded in this—peoples' energy impact on a community, has been notoriously invisible. Researchers have extensively pointed to the importance of mobilizing people to take action, in conjunction with institutional actions, in adopting less energy-intensive lifestyles (Bomberg and McEwen 2012; Schot et al. 2016). While various dashboards and web-based platforms have visualized incoming public community energy data, these platforms have been primarily developed for use by entities in the real-estate sector or those with the capital and resources to make sense of the data (Gulbinas and Jain 2016; Kontokosta and Tull 2015). It is important to understand how citizens respond to having access to this data and explore potential uses of this information. To contribute to this effort, this study investigates transforming community-scale energy data, which is becoming increasingly available across communities, into a form that is more usable, engaging, and actionable to community members.

The developed application is a mobile application integrating Georgia Tech's building energy consumption and production data, presented in the context of Georgia Tech's community-wide current performance and future goals. The design and validation of the developed systems contributes to the literature in several ways. First, it builds on previous building energy feedback systems by expanding energy feedback to the community scale. Several prior energy feedback studies have suggested the potential of community-scale energy feedback to further engage users (Geelen et al. 2013; Pierce and Paulos 2012b), and several studies have integrated elements of community-based communication and neighborhood maps within energy feedback (Burchell et al. 2016;

Gupta et al. 2017). The documented framework and open-source code for this application builds on this work by drawing on successful features documented in the energy feedback literature, integrated these functionalities into the application, and broadening the data represented to the community scale. The new community-based features which add context to the energy information include integration of campus energy sources, inclusion of community goals, representation of community-wide performance, and connection of energy data to the physical infrastructure of a community.

Secondly, thinking aloud procedures were applied in the evaluation of the application, which have been used effectively for prototype evaluation in the field of human-computer interaction for decades (Nielsen 1994). Engaging potential prospective users during the design stage of an interface is critical yet often overlooked step in the development of energy technologies. It has been well documented how interfaces designed for the designer, rather than the intended user, can substantially impact of effectiveness of the technology during implementation (Geelen et al. 2013; Skjølsvold and Lindkvist 2015). In response to this gap, the evaluation procedure conducted engaged users with a prototype of a relatively novel interface, community energy feedback, to gather perspectives of the interface and the community energy data in general.

The thinking aloud evaluation procedure identified people’s initial reactions to having access to community energy data and prompted several areas for future research. As a whole, the observation and survey results indicated there was a high interest among most participants in having access to community energy data across different geographic scales (i.e., the Georgia Tech campus, the neighborhood, and the city level). Thus, it would be fruitful for future studies to implement and study community scale energy feedback systems beyond the university campus scale. The results of this study provide an important basis for future research to extend this application to new scales and test user engagement in various environments. While the developed community-scale energy feedback system has the potential to improve community understanding of and engagement with energy issues, the application has yet to incorporate of action-oriented elements. Directly linking information to action is important for behavior change (Jensen 2010), and future areas of research expanding the development of these systems should work to include action elements. Elements could be personalized to one's own community, such as with specific upcoming energy events, as well as support social engagement and connection.

5.3 Dependencies between infrastructure and solar adoption

In a final study, I took a broader approach in examining energy trends at the urban-scale. Specifically, I investigated the relationship between renewable energy adoption in cities and the design of city infrastructure. Many urban environments have committed to clean energy goals, with solar PV often presented as a viable option in helping achieve a large part of this vision (SierraClub 2019). Rooftop solar PV is reliant on building infrastructure to be installed and perform efficiently; shading from densification, new developments, and evolving urban designs have potential to restrict solar PV adoption through limiting existing building rooftop access to solar and increasing the risk of a solar system being obstructed in the future. While solar PV adoption analyses in the public policy and social science domains have studied a myriad of factors that are theorized to affect solar adoption (e.g., financial incentives, adopter attitudes, peer effects), built infrastructure is often omitted in such analyses, or is characterized in such a way that is difficult to draw meaning from the results. This study extends solar adoption literature by expanding on the theoretical basis for including built infrastructure as a factor associated with solar adoption, characterizing the built environment with a more direct metric, and quantifying the relationship between this metric and solar adoption.

The paper expands on the definition of the built environment in solar adoption analyses and theoretical foundations for including this metric. I discuss the limitations of using housing density to represent the built environment by differentiating between two foundationally divergent factors that are both embedded with housing density metrics: solar rights issues and rooftop solar suitability. Inspired by the extensive work conducted in computational engineering and architecture domains to quantify city rooftop solar suitability (Lobaccaro et al. 2019; Sigrin and Mooney 2018), I introduce rooftop solar suitability as a metric in solar PV adoption modeling. The results show that rooftop solar suitability is significantly associated with solar PV adoption across metropolitan regions in the Southeast US, while controlling for other variables commonly associated with solar adoption. Current solar adoption analyses tend to focus on the importance of social or financial factors in influencing solar PV adoption (Schelly 2014). This finding suggests that rooftop solar suitability, or more broadly, urban form, may also be an important factor to consider in solar PV modeling and is currently overlooked.

Scholars have brought attention to in recent years the importance of studying interactions and unforeseen outcomes between the built infrastructure systems and available climate change solutions (Creutzig et al. 2016). This study seeks to contribute to this call by examining dependencies between renewable energy infrastructure (i.e., solar PV systems) and the urban forms they depend upon (i.e., rooftop solar suitability). This is a first step in understanding the relationship between urban forms and one major solution cities are looking towards as they

The results of this study provide a basis for a multitude of research opportunities. The findings establish an interesting link between rooftop solar suitability and solar adoption. As more data is released on rooftop solar suitability (Sigrin and Mooney 2018; Yu et al. 2018) or as existing algorithms are applied on high-resolution historic satellite imagery (Google 2020), additional analyses can be conducted that include intra-census tract temporal variation in solar adoption and rooftop solar suitability. Such an approach can enable causal links between urban form and solar adoption to be examined. Additionally, there are many alternative ways of representing urban form beyond rooftop solar suitability. Future analyses could define urban form using alternative metrics, which may provide additional insights regarding the relationship between urban form and renewable energy adoption. Finally, a natural avenue for future research from a policy standpoint is to examine the built environment policies that impact rooftop solar suitability. Ultimately, variations in rooftop solar suitability will depend upon built environment policy; the link between rooftop solar suitability and solar PV adoption provided in this study gives impetus for future research to study the policies that have the greatest impact on rooftop solar suitability.

CHAPTER 6. CONCLUSION

City and community level initiatives offer promising contexts for sustainability agendas. As local initiatives aimed at reducing carbon emissions rise, and sensors become more affordable and ubiquitous across community infrastructure, the potential of data to aid decision making has become a prominent focus, particularly in the building sector. The fields of urban informatics, energy data analytics, and computational sciences are growing immensely with the influx of new ways to collect data on built infrastructure and people. Across my dissertation research, each study contains a new approach for transforming an emerging source of data into more understandable, engaging, and actionable information for specific stakeholders. These studies draw on principles from data science, user-centered design, and public policy to develop new directions for fields aiming to reduce buildings' reliance on fossil-fuels and create sustainable built infrastructure. Continued connections drawn between these areas are likely to foster unique insights, encourage participation and better understanding of energy issues, and pave the way for communities to attain a low-carbon future.

APPENDIX A. ADDITIONAL FIGURES AND TABLES

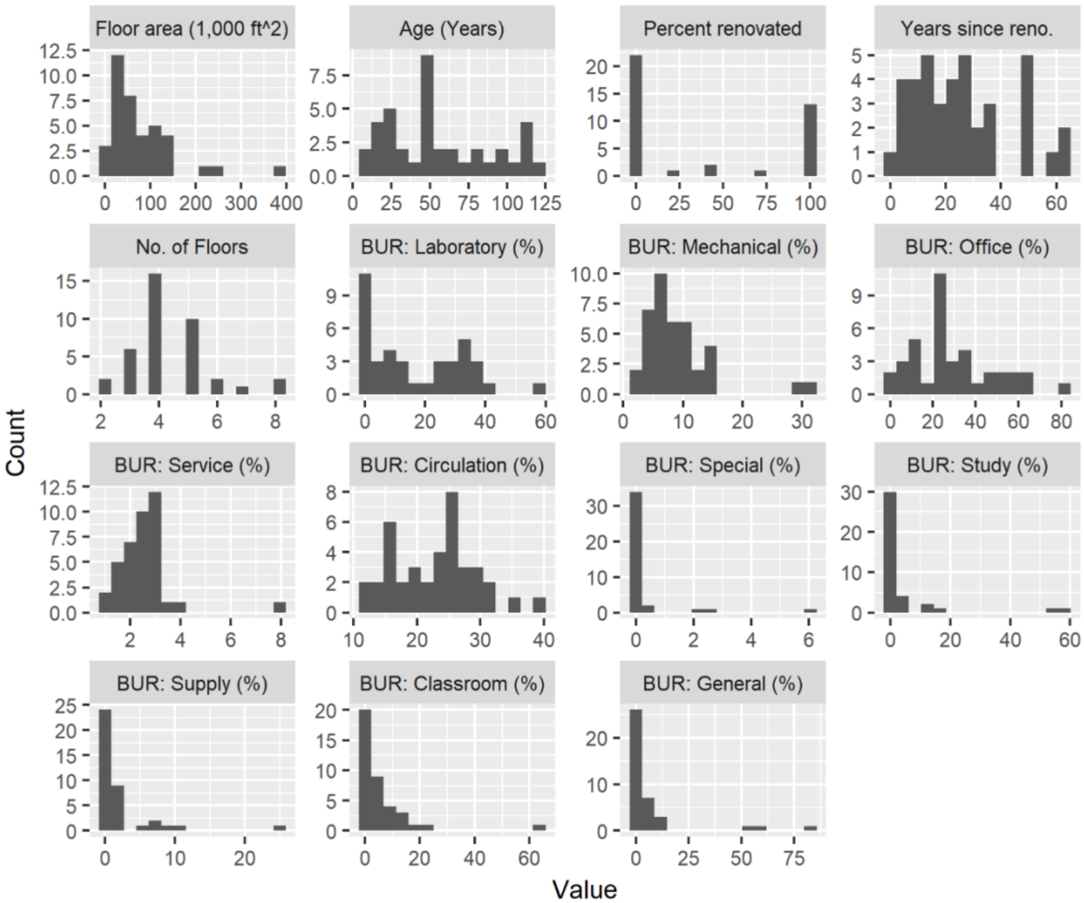


Figure 16 – Explanatory Variable Distributions.

Each histogram displays the distribution of each explanatory variable input into the regression models.

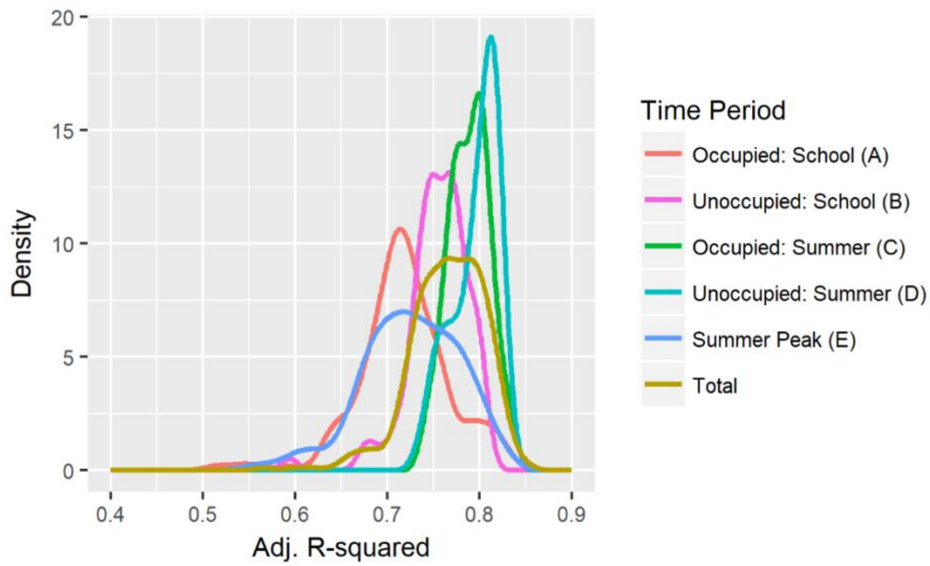
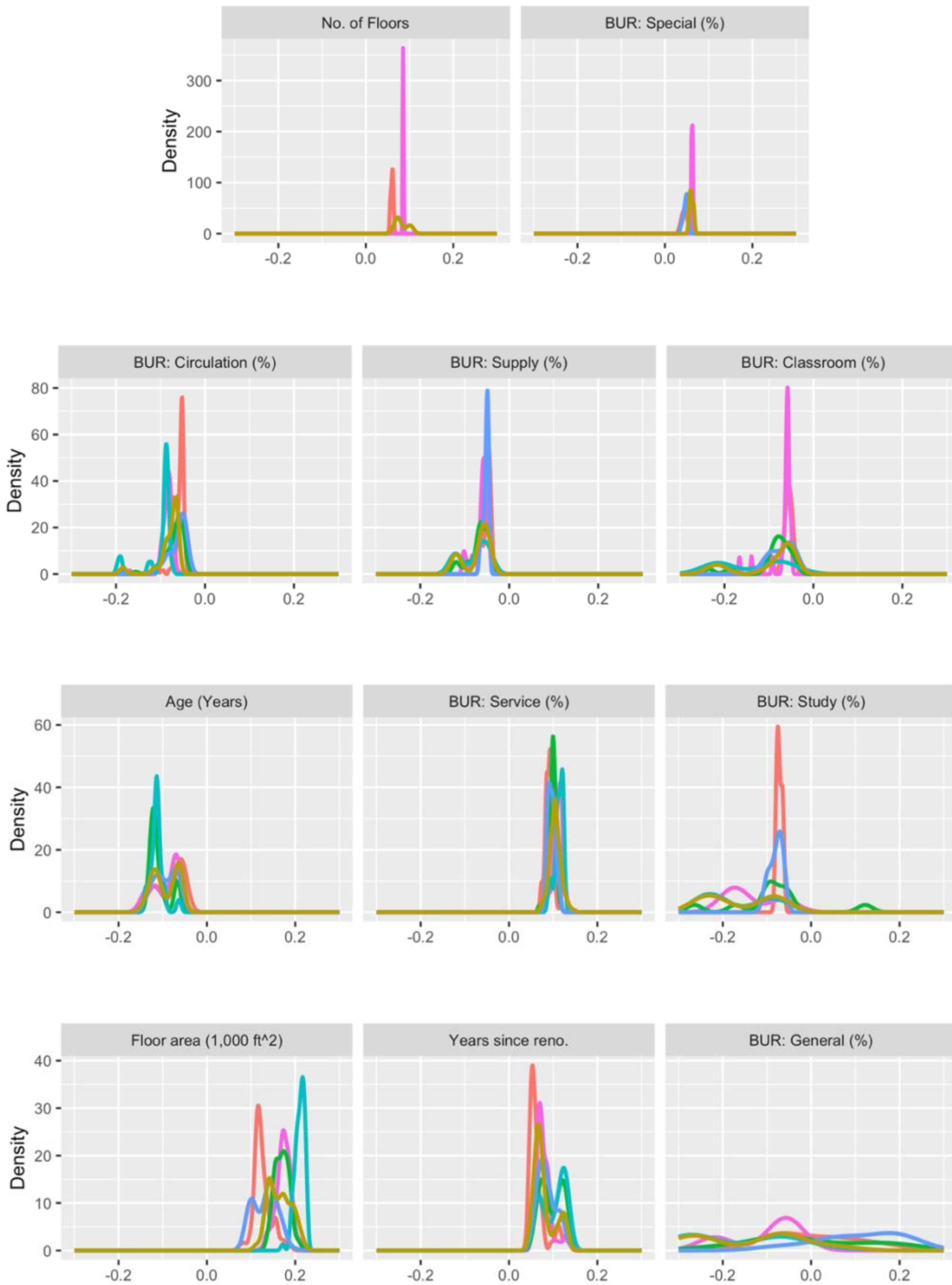


Figure 17 – Adjusted R-squared Values for Regression Models across all Temporal Periods.

Each line represents the distribution of the regression model fit for a temporal period.


No. of Floors
BUR: Special (%)
BUR: Circulation (%)
BUR: Supply (%)
BUR: Classroom (%)
Age (Years)
BUR: Service (%)
BUR: Study (%)
Floor area (1,000 ft^2)
Years since reno.
BUR: General (%)
Density
300
200
100
80
60
40
20
0
-0.2
0.0
0.2

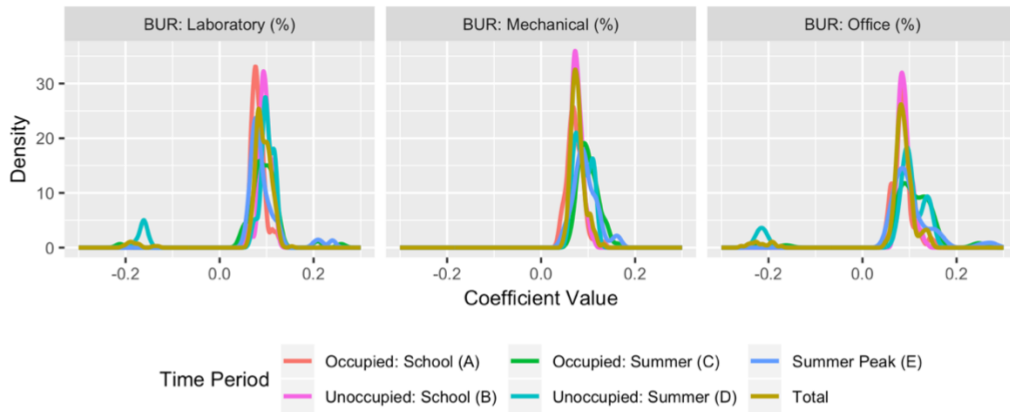


Figure 18 – Significant Regression Coefficient Distributions for All Regression Model Independent Variables.

For each density plot, a line represents the distribution of regression coefficients used to benchmark energy consumption for each day during a temporal period in a probability density function. Only significant regression coefficients are included in the density plots. Each row is grouped by comparable density probability ranges for visual purposes.

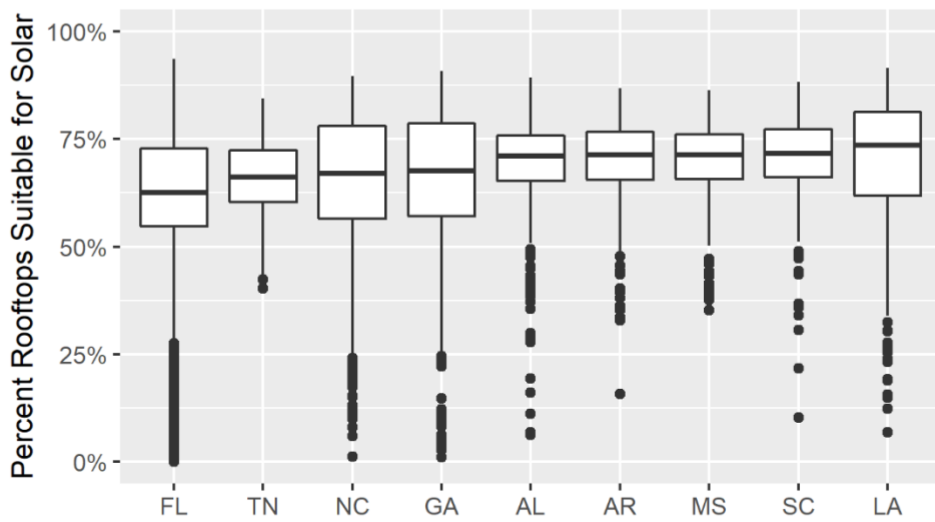


Figure 19 – Distribution of rooftop solar suitability by state

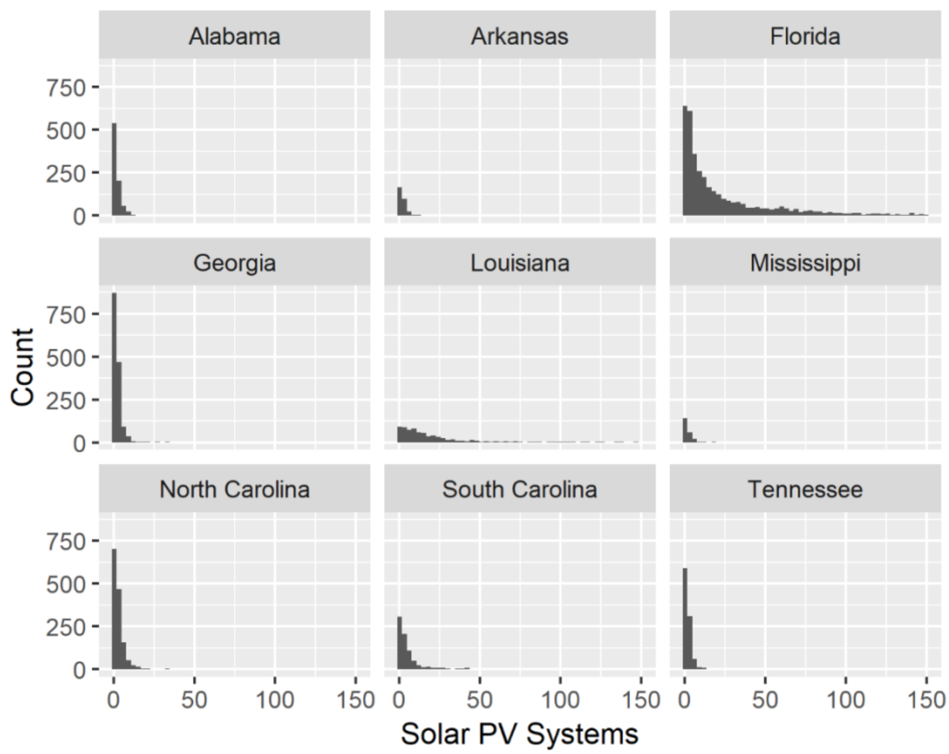


Figure 20 – Count of solar PV systems installed per census tract, by state

Table 15 – Model results for solar system count per number of residences in a census tract

	Raw covariates	Standardized covariates
Solar radiation	2.118*** (0.065)	0.672*** (0.02)
Solar suitability	**4.679*** (0.154)**	**0.673*** (0.022)**
Log(income)	0.671*** (0.069)	0.288*** (0.03)
Sqrt(education)	0.908*** (0.158)	0.148*** (0.026)
Employment	0.762* (0.305)	0.045* (0.018)
Political affiliation	1.249*** (0.136)	0.198*** (0.021)
Race- white	-0.628*** (0.075)	-0.168*** (0.02)
Log(housing density)	0.09*** (0.012)	0.145*** (0.019)
Rented housing	1.608*** (0.247)	0.182*** (0.028)
Electricity rate	4.701*** (1.352)	0.053*** (0.015)
Net metering policy[a]	0.677*** (0.054)	0.677*** (0.054)
Property tax policy[a]	1.21*** (0.036)	1.21*** (0.036)
Tax credit policy[a]	-0.067 (0.042)	-0.067 (0.042)

*Statistical significance: ***(p < 0.001), **(p < 0.01), *(p < 0.05);* [a] *Variable is binary*

Table 16 – Standardized coefficient results for models without policy variables, with individual policy variables, and with aggregated policy variables

	(1) Negative Binomial	(2) Negative Binomial	(3) Negative Binomial	(4) Negative Binomial	(5) Negative Binomial
Intercept	2.089*** (0.014)	1.969*** (0.014)	1.918*** (0.013)	0.506*** (0.045)	1.897*** (0.013)
Log(population)	0.377*** (0.014)	0.386*** (0.014)	0.406*** (0.014)	0.381*** (0.013)	0.395*** (0.014)
Solar radiation	1.046*** (0.014)	1.01*** (0.014)	0.97*** (0.015)	0.704*** (0.02)	0.938*** (0.015)
Solar suitability	**0.273*** (0.014)**	**0.611*** (0.019)**	**0.716*** (0.023)**	**0.666*** (0.022)**	**0.708*** (0.023)**
Log(income)		0.084** (0.026)	0.166*** (0.031)	0.201*** (0.029)	0.19*** (0.031)
Sqrt(education)		0.224*** (0.025)	0.154*** (0.026)	0.206*** (0.025)	0.149*** (0.026)
Employment		0.079*** (0.019)	0.051** (0.018)	0.065*** (0.018)	0.064*** (0.018)
Political affiliation		0.425*** (0.02)	0.371*** (0.022)	0.164*** (0.021)	0.293*** (0.022)
Race- white		-0.062** (0.02)	-0.021 (0.02)	-0.144*** (0.02)	-0.053** (0.02)
Log(housing density)			0.309*** (0.019)	0.227*** (0.019)	0.297*** (0.019)
Rented housing			0.052 (0.029)	0.114*** (0.028)	0.087** (0.029)
Electricity rate			-0.205*** (0.014)	0.055*** (0.015)	-0.156*** (0.014)
Net metering policy[a]				0.668*** (0.054)	
Property tax policy[a]				1.197*** (0.035)	
Tax credit policy[a]				-0.076 (0.042)	
Policy total					0.249*** (0.015)

*Statistical significance: ***(p < 0.001), **(p < 0.01), *(p < 0.05); [a] Variable is binary*

REFERENCES

Abrahamse, W., Steg, L., Vlek, C., and Rothengatter, T. (2005). “A review of intervention studies aimed at household energy conservation.” *Journal of Environmental Psychology*, 25(3), 273–291.

Abras, C., Maloney-krichmar, D., and Preece, J. (2004). “User Centered Design.” *Design*, Thousand Oaks: Sage Publications, 37(4), 1–14.

ACS. (2016). “2011-2015 ACS 5-year Estimates.” *US Census*, <https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2015/5-year.html> (Mar. 12, 2019).

AndrewProjDent. (2018). “ARKit + CoreLocation.” *Github*, <https://github.com/ProjectDent/ARKit-CoreLocation>.

“Apple Developer.” (2018). *Apple*, <https://developer.apple.com/>.

Attard, J., Orlandi, F., Scerri, S., and Auer, S. (2015). “A systematic review of open government data initiatives.” *Government Information Quarterly*, Elsevier Inc., 32(4), 399–418.

Azar, E., and Al Ansari, H. (2017). “Framework to investigate energy conservation motivation and actions of building occupants: The case of a green campus in Abu Dhabi, UAE.” *Applied Energy*, Elsevier Ltd, 190, 563–573.

Baxter, R., Hastings, N., Law, A., and Glass, E. J. (2011). “A Theory of Smart Cities.” *Proceedings of the 55th Annual Meeting of the International Society for the Systems Sciences*, 1–15.

Bollinger, B., and Gillingham, K. (2012). “Peer Effects in the Diffusion of Solar Photovoltaic Panels.” *Marketing Science*, 31(6), 900–912.

Bomberg, E., and McEwen, N. (2012). “Mobilizing community energy.” *Energy Policy*, Elsevier, 51, 435–444.

Bonino, D., Corno, F., and De Russis, L. (2012). “Home energy consumption feedback: A user survey." *Energy and Buildings*, Elsevier B.V., 47, 383–393.

Borgstein, E. H., Lamberts, R., and Hensen, J. L. M. (2016). "Evaluating energy performance in non-domestic buildings: A review." *Energy and Buildings*, Elsevier B.V., 128, 734–755.

Bretz, F., Hothorn, T., and Westfall, P. (2011). *Multiple Comparisons Using R*. Chapman and Hall/CRC.

Bronin, S. (2009). "Solar Rights." *Boston University Law Review*, 89(4), 1217–1266.

Buck, J., and Young, D. (2007). "The potential for energy efficiency gains in the Canadian commercial building sector: A stochastic frontier study." *Energy*, 32(9), 1769–1780.

BuildingRating. (2019). "Sharing Transparency for a More Efficient Future." *BuildingRating*, <https://www.buildingrating.org/> (Jul. 30, 2017).

BuildSmart DC. (2017). "Buildings." <http://www.buildsmartdc.com/> (Jul. 30, 2017).

Burchell, K., Rettie, R., and Roberts, T. C. (2016). "Householder engagement with energy consumption feedback: The role of community action and communications." *Energy Policy*, Elsevier Ltd, 88, 168–177.

Burke, L., Reytar, K., Spalding, M., and Perry, A. (2011). *Reefs at Risk Revisited*. *Defenders*.

Burman, E., Hong, S. M., Paterson, G., Kimpian, J., and Mumovic, D. (2014). "A comparative study of benchmarking approaches for non-domestic buildings: Part 2 - Bottom-up approach." *International Journal of Sustainable Built Environment*, The Gulf Organisation for Research and Development, 3(2), 247–261.

C40 Cities. (2020). "C40 Cities." *C40 Cities*, <https://www.c40.org/>.

Camagni, R., Cristina, M., and Rigamonti, P. (2002). "Urban mobility and urban form: the social and environmental costs of different patterns of urban expansion." *Ecological Economics*, 40, 199–216.

CBECS. (2015). "A Look at the U.S. Commercial Building Stock: Results from EIA's

2012 Commercial Buildings Energy Consumption Survey.” *U.S. Energy Information Administration*, <https://www.eia.gov/consumption/commercial/reports/2012/buildstock/>.

Cheng, V., Steemers, K., Montavon, M., and Compagnon, R. (2006). “Urban Form, Density and Solar Potential.” *The 23rd Conference on Passive and Low Energy Architecture*.

Chester, M. V., and Allenby, B. (2019). “Toward adaptive infrastructure: flexibility and agility in a non-stationarity age.” *Sustainable and Resilient Infrastructure*, Taylor & Francis, 4(4), 173–191.

“Chicago Energy Benchmarking.” (2018). *City Energy*, <http://cityenergyproject.github.io/chicago/>.

Chourabi, H., Nam, T., Walker, S., Gil-Garcia, J. R., Mellouli, S., Nahon, K., Pardo, T. A., and Scholl, H. J. (2011). “Understanding smart cities: An integrative framework.” *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2289–2297.

Chung, W. (2011). “Review of building energy-use performance benchmarking methodologies.” *Applied Energy*, Elsevier Ltd, 88(5), 1470–1479.

Chung, W. (2012). “Using the fuzzy linear regression method to benchmark the energy efficiency of commercial buildings.” *Applied Energy*, Elsevier Ltd, 95, 45–49.

Chung, W., Hui, Y. V., and Lam, Y. M. (2006). “Benchmarking the energy efficiency of commercial buildings.” *Applied Energy*, 83(1), 1–14.

Cisco. (2020). “Cities and Communities.” *Cisco*, <https://www.cisco.com/c/en/us/solutions/industries/smart-connected-communities.html>.

Creutzig, F., Agoston, P., Minx, J. C., Canadell, J. G., Andrew, R. M., Quéré, C. Le, Peters, G. P., Sharifi, A., Yamagata, Y., and Dhakal, S. (2016). “Urban infrastructure choices structure climate solutions.” *Nature Climate Change*, Nature Publishing Group, 6(12), 1054–1056.

D’Aniello, F., Sorrentino, M., Rizzo, G., Trifirò, A., and Bedogni, F. (2018). “Introducing innovative energy performance metrics for high-level monitoring and diagnosis of telecommunication sites." *Applied Thermal Engineering*, 137(March), 277–287.

Dharshing, S. (2017). "Household dynamics of technology adoption: A spatial econometric analysis of residential solar photovoltaic (PV) systems in Germany." *Energy Research and Social Science*, Elsevier Ltd, 23, 113–124.

Dixon, G. N., Deline, M. B., McComas, K., Chambliss, L., and Hoffmann, M. (2015). "Using Comparative Feedback to Influence Workplace Energy Conservation: A Case Study of a University Campaign." *Environment and Behavior*, 47(6), 667–693.

DSIRE. (2019). "Database of State Incentives for Renewables & Efficiency." <http://www.dsireusa.org/>.

EIA. (2018). "How many smart meters are installed in the United States, and who has them?" *U.S. Energy Information Administration*, <https://www.eia.gov/tools/faqs/faq.php?id=108&t=3> (Mar. 29, 2019).

ElYamany, H. F., Seewald, L., Grolinger, K., Higashino, W. A., Capretz, M. A. M., ElYamany, H. F., Higashino, W. A., Capretz, M. A. M., Seewald, L., Grolinger, K., Higashino, W. A., and Capretz, M. A. M. (2017). "Energy slices: benchmarking with time slicing." *Energy Efficiency*, Energy Efficiency, 11(2), 521–538.

Energy Information Administration. (2019). "How much energy is consumed in U.S. residential and commercial buildings?" *US EIA*, <https://www.eia.gov/tools/faqs/faq.php?id=86&t=1>.

EuropeanCommission. (2019). "Energy performance of buildings."

Fischer, C. (2008). "Feedback on household electricity consumption: a tool for saving energy?" *Energy Efficiency*, 1(1), 79–104.

Francisco, A., Mohammadi, N., and Taylor, J. E. (2018a). "Evaluating Temporal Shifts in City Scale Building Energy Benchmarks." *Proceeding of Construction Research Congress 2018*, 450–460.

Francisco, A., Mohammadi, N., and Taylor, J. E. (2020). "Smart City Digital Twin-Enabled Energy Management: Toward Real-Time Urban Building Energy Benchmarking." *Journal of Management in Engineering*, 36(2).

Francisco, A., and Taylor, J. E. (2019a). “Understanding citizen perspectives on open urban energy data through the development and testing of a community energy feedback system.” *Applied Energy*, 256.

Francisco, A., and Taylor, J. E. (2019b). “Designing community-scale energy feedback.” *Proceedings of the 10th International Conference on Applied Energy*, Elsevier B.V., 158, 4178–4183.

Francisco, A., Truong, H., Khosrowpour, A., Taylor, J. E., and Mohammadi, N. (2018b). “Occupant perceptions of building information model-based energy visualizations in eco-feedback systems.” *Applied Energy*, Elsevier, 221(December 2017), 220–228.

Gagnon, P., Margolis, R., Melius, J., Phillips, C., and Elmore, R. (2016). *Rooftop Solar Photovoltaic Technical Potential in the United States: A Detailed Assessment*. *National Renewable Energy Laboratory*.

Geelen, D., Reinders, A., and Keyson, D. (2013). “Empowering the end-user in smart grids: Recommendations for the design of products and services.” *Energy Policy*, Elsevier, 61, 151–161.

Gillingham, K., Harding, M., and Rapson, D. (2012). “Split incentives in residential energy consumption.” *Energy Journal*, 33(2), 37–62.

Gindi, D., and Jahoda, P. (2016). “Charts.” *Github*, <https://github.com/danielgindi/Charts>.

Google. (2020). “Project Sunroof.” *Google*, <https://www.google.com/get/sunroof> (Feb. 2, 2020).

Graziano, M., Fiaschetti, M., and Atkinson-Palombo, C. (2019). “Peer effects in the adoption of solar energy technologies in the United States: An urban case study.” *Energy Research and Social Science*, Elsevier, 48(March 2017), 75–84.

Graziano, M., and Gillingham, K. (2015). “Spatial patterns of solar photovoltaic system adoption: The influence of neighbors and the built environment.” *Journal of Economic Geography*, 15(4), 815–839.

Gul, M. S., and Patidar, S. (2015). “Understanding the energy consumption and occupancy of a multi-purpose academic building.” *Energy and Buildings*, Elsevier B.V., 87, 155–

165.

Gulbinas, R., and Jain, R. (2016). “Towards the development of a visual data exploration tool to augment decision-making in urban building energy efficiency programs.” *16th International Conference on Computing in Civil and Building Engineering*.

Gulbinas, R., Jain, R. K., and Taylor, J. E. (2014). “BizWatts: A modular socio-technical energy management system for empowering commercial building occupants to conserve energy.” *Applied Energy*, Elsevier Ltd, 136, 1076–1084.

Gulbinas, R., Khosrowpour, A., and Taylor, J. (2015). “Segmentation and Classification of Commercial Building Occupants by Energy-Use Efficiency and Predictability.” *IEEE Transactions on Smart Grid*, 6(3), 1414–1424.

Gupta, R., Barnfield, L., and Gregg, M. (2017). “Exploring innovative community and household energy feedback approaches.” *Building Research and Information*, Taylor & Francis, 1–16.

Gupta, S. K. S., Mukherjee, T., Varsamopoulos, G., and Banerjee, A. (2011). “Research directions in energy-sustainable cyberphysical systems.” *Sustainable Computing: Informatics and Systems*, Elsevier Inc., 1(1), 57–74.

Gupta, V. (1997). “Thermal Efficiency of Building Clusters: an Index for Non Air-Conditioned Buildings in Hot Climates.” *Energy and Urban Built Form*, (1974), 133–145.

Gurstein, M. (2011). “Open data: Empowering the empowered or effective data use for everyone?” *First Monday*, 16(2), 1–5.

Hachem, C., Athienitis, A., and Fazio, P. (2011). “Parametric investigation of geometric form effects on solar potential of housing units.” *Solar Energy*, Elsevier Ltd, 85(9), 1864–1877.

Hastak, M., and Koo, C. (2016). “Theory of an Intelligent Planning Unit for the Complex Built Environment.” *Journal of Management in Engineering*, 33(3), 04016046.

Hilbe, J. M. (2011). *Negative Binomial Regression*. Cambridge University Press.

Hollands, R. G. (2015). "Critical interventions into the corporate smart city." *Cambridge Journal of Regions, Economy and Society*, 8(1), 61–77.

Holm, S. (1979). "A Simple Sequentially Rejective Multiple Test Procedure." *Scandinavian Journal of Statistics*, 6(2), 65–70.

Hong, T., Kim, H., and Kwak, T. (2012). "Energy-saving techniques for reducing CO 2 emissions in elementary schools." *Journal of Management in Engineering*, 28(1), 39–50.

Hsu, J. H. Y. (2018). "Predictors for adoption of local solar approval processes and impact on residential solar installations in California cities." *Energy Policy*, Elsevier Ltd, 117(December 2017), 463–472.

IBM. (2020). "Government modernization is critical to serve citizens in the digital era." *IBM*, <https://www.ibm.com/industries/government>.

IMT. (2015). "Chicago's Large Buildings Could Save $184 Million Annually." *Institute for Market Transformation*, <https://www.imt.org/boosting-energy-efficiency-in-chicagos-large-buildings-could-save-up-to-184/> (May 1, 2018).

Institute for Market Transformation. (2019). "Comparison of U.S. Commercial Building Energy Benchmarking and Transparency Policies." *Institute for Market Transformation*, <https://www.imt.org/resources/comparison-of-commercial-building-benchmarking-policies/>.

Institute for Market Transformation. (2020). *Energy benchmarking and transparency*.

IPCC. (2007). *Working group III contribution to the fourth assessment report of the Intergovernmental Panel on Climate Change*. (B. Metz, O. Davidson, P. Bosch, R. Dave, and L. Meyer, eds.), Cambridge University Press.

Jain, R. K., Taylor, J. E., and Peschiera, G. (2012). "Assessing eco-feedback interface usage and design to drive energy efficiency in buildings." *Energy and Buildings*, Elsevier B.V., 48, 8–17.

Jensen, B. B. (2010). "Knowledge , Action and Pro- environmental Behaviour Knowledge , Action and Pro-environmental Behaviour." *Environmental Education Research*, 8(3), 325–334.

Jindal, A., Dua, A., Kaur, K., Singh, M., Kumar, N., and Mishra, S. (2016). "Decision Tree and SVM-Based Data Analytics for Theft Detection in Smart Grid." *IEEE Transactions on Industrial Informatics*, IEEE, 12(3), 1005–1016.

Joseph, S., and Erakkath Abdu, J. (2018). "Real-time retail price determination in smart grid from real-time load profiles." *International Transactions on Electrical Energy Systems*, 28(3), 1–11.

Kavousian, A., and Rajagopal, R. (2014). "Data-Driven Benchmarking of Building Energy Efficiency Utilizing Statistical Frontier Models." *Journal of Computing in Civil Engineering*, 28(1), 79–88.

Kettles, C. M. (2008). *A Comprehensive Review of Solar Access Law in the United States*.

Khan, Z., Anjum, A., Soomro, K., and Tahir, M. A. (2015). "Towards cloud based big data analytics for smart future cities." *Journal of Cloud Computing: Advances, Systems, and Applications*, 4(2).

Kinney, S., and Piette, M. A. (2002). "Development of a California Commercial Building Energy Benchmarking Database." *2002 ACEEE Summer Study on Energy Efficiency in Buildings*, 109–120.

Kitchin, R., Lauriault, T. P., and McArdle, G. (2015). "Knowing and governing cities through urban indicators, city benchmarking and real-time dashboards." *Regional Studies, Regional Science*, Routledge, 2(1), 6–28.

Klein-Banai, C., and Theis, T. L. (2011). "An urban university's ecological footprint and the effect of climate change." *Ecological Indicators*, Elsevier Ltd, 11(3), 857–860.

Koenker, R. (2019). *Quantile Regression*.

Koenker, R., and Hallock, K. (2001). "Quantile regression." *Journal of Economic Perspectives*, 15(4), 143–156.

Kontokosta, C. E. (2013). "Energy disclosure, market behavior, and the building data ecosystem." *Annals of the New York Academy of Sciences*, 1295(1), 34–43.

Kontokosta, C. E., and Tull, C. (2015). "EnergyViz: Web-Based Eco-Visualization of

Urban Energy Use from Building Benchmarking Data." *Proceedings of the International Conference on Computing in Civil and Building Engineering*, (March), 1405–1412.

Koo, C., Kim, J., Jeong, K., Hong, T., and Lee, M. (2015). "A review on sustainable construction management strategies for monitoring, diagnosing, and retrofitting the building's dynamic energy performance: Focused on the operation and maintenance phase." *Applied Energy*, Elsevier Ltd, 155, 671–707.

Koomey, J., and Brown, R. E. (2002). "The role of building technologies in reducing and controlling peak electricity demand." *Lawrence Berkeley National Laboratory*, (September).

Kwan, C. L. (2012). "Influence of local environmental, social, economic and political variables on the spatial distribution of residential solar PV arrays across the United States." *Energy Policy*, Elsevier, 47, 332–344.

"LA Energy Atlas." (2018). *UCLA California Center for Sustainabile Communities*, <http://www.energyatlas.ucla.edu/>.

Lasco Crago, C., and Chernyakhovskiy, I. (2017). "Are policy incentives for solar power effective? Evidence from residential installations in the Northeast." *Journal of Environmental Economics and Management*, Elsevier, 81, 132–151.

Li, H., and Yi, H. (2014). "Multilevel governance and deployment of solar PV panels in U.S. cities." *Energy Policy*, Elsevier, 69, 19–27.

Li, Z., Han, Y., and Xu, P. (2014). "Methods for benchmarking building energy consumption against its past or intended performance: An overview." *Applied Energy*, Elsevier Ltd, 124, 325–334.

Lobaccaro, G., Carlucci, S., Croce, S., Paparella, R., and Finocchiaro, L. (2017). "Boosting solar accessibility and potential of urban districts in the Nordic climate: A case study in Trondheim." *Solar Energy*, Elsevier Ltd, 149, 347–369.

Lobaccaro, G., Lisowska, M. M., Saretta, E., Bonomo, P., and Frontini, F. (2019). "A methodological analysis approach to assess solar energy potential at the neighborhood scale." *Energies*, 12(18).

Lukanov, B. R., and Krieger, E. M. (2019). “Distributed solar and environmental justice: Exploring the demographic and socio-economic trends of residential PV adoption in California.” *Energy Policy*, Elsevier Ltd, 134(April), 110935.

Mann, M. E., and Emanuel, K. A. (2006). “Atlantic Hurricane trends linked to climate change.” *Eos*, 87(24), 233–241.

Masoso, O. T., and Grobler, L. J. (2010). “The dark side of occupants’ behaviour on building energy use.” *Energy and Buildings*, 42(2), 173–177.

Matisoff, D. C., and Johnson, E. P. (2017). “The comparative effectiveness of residential solar incentives.” *Energy Policy*, 108(May), 44–54.

McAdam, J. (2012). *Climate change, forced migration, and international law*. Oxford University Press.

Mildenberger, M., Howe, P. D., and Miljanich, C. (2019). “Households with solar installations are ideologically diverse and more politically active.” *Nature Energy*, Springer US.

Mohammadi, N., and Taylor, J. E. (2017). “Smart City Digital Twins.” *Proceedings of the 2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, IEEE, 1–5.

Mordor Intelligence. (2019). *Big Data Analytics Market in the Energy Sector- Growth, Trends, and Forecast*.

NASA. (2008). “NASA Surface Meteorology and Solar Energy.” *NASA*, <https://ntrs.nasa.gov/search.jsp?R=20080012141> (Mar. 12, 2019).

Neufeld, J. (1987). “Price Discrimination and the Adoption of the Electricity Demand Charge.” *The Journal of Economic History*, 47(3), 693–709.

Nielsen, J. (1994). “Estimating the number of subjects needed for a thinking aloud test.” *International Journal of Human - Computer Studies*.

Nolan, J. M., Schultz, P. W., Cialdini, R. B., Goldstein, N. J., and Griskevicius, V. (2008). “Normative social influence is underdetected.” *Personality & social psychology bulletin*, 34(7), 913–23.

NREL. (2017). "U.S. Electric Utility Companies and Rates: Look-up by Zipcode." *NREL*, <https://openei.org/doe-opendata/dataset/u-s-electric-utility-companies-and-rates-look-up-by-zipcode-2017> (Jan. 3, 2020).

O'Dwyer, E., Pan, I., Acha, S., and Shah, N. (2019). "Smart energy systems for sustainable smart cities: Current developments, trends and future directions." *Applied Energy*, Elsevier, 237(October 2018), 581–597.

O'Hara, R. B., and Kotze, D. J. (2010). "Do not log-transform count data." *Methods in Ecology and Evolution*, 1(2), 118–122.

Ornetzeder, M., and Rohracher, H. (2006). "User-led innovations and participation processes: Lessons from sustainable energy technologies." *Energy Policy*, 34(2 SPEC. ISS.), 138–150.

Ornetzeder, M., and Rohracher, H. (2013). "Of solar collectors, wind power, and car sharing: Comparing and understanding successful cases of grassroots innovations." *Global Environmental Change*, Elsevier Ltd, 23(5), 856–867.

Oster, E. (2019). "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business and Economic Statistics*, Taylor & Francis, 37(2), 187–204.

Palmer, K., and Walls, M. (2015). "Can Benchmarking and Disclosure Laws Provide Incentives for Energy Efficiency Improvements in Buildings ?" Resources For the Future, 1–29.

Palmer, K., and Walls, M. (2016). "Using information to close the energy efficiency gap: a review of benchmarking and disclosure ordinances." *Energy Efficiency*, Energy Efficiency, 1–19.

Park, H. S., Lee, M., Kang, H., Hong, T., and Jeong, J. (2016). "Development of a new energy benchmark for improving the operational rating system of office buildings using various data-mining techniques." *Applied Energy*, Elsevier Ltd, 173, 225–237.

Peacock, A. D., Chaney, J., Goldbach, K., Walker, G., Tuohy, P., Santonja, S., Todoli, D., and Owens, E. H. (2017). "Co-designing the next generation of home energy management systems with lead-users." *Applied Ergonomics*, Elsevier Ltd, 60, 194–206.

Pérez-Lombard, L., Ortiz, J., González, R., and Maestre, I. R. (2009). “A review of benchmarking, rating and labelling concepts within the framework of building energy certification schemes.” *Energy and Buildings*, 41(3), 272–278.

Petkov, P., Foth, M., Köbler, F., and Krcmar, H. (2011). “Motivating Domestic Energy Conservation through Comparative, Community-Based Feedback in Mobile and Social Media Queensland University of Technology Chair for Information Systems.” *Proceedings of the 5th International Conference on Communities and Technologies - C&T ’11*, 21–30.

Pfeiffer, S. (1982). “Ancient Lights: Legal Protection of Access to Solar Energy.” *American Bar Association Journal*, 68(3), 288–291.

Pierce, J., and Paulos, E. (2012a). “Beyond energy monitors: Interaction, Energy, and Emergining Energy Systems.” *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, 665–674.

Pierce, J., and Paulos, E. (2012b). “Beyond energy monitors: interaction, energy, and emerging energy systems.” *CHI ’12: Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, 1–10.

Pierce, P., and Andersson, B. (2017). “Challenges with smart cities initiatives – A municipal decision makers ’ perspective.” 2804–2813.

Poovendran, R. (2010). “Cyber-physical systems: Close encounters between two parallel worlds.” *Proceedings of the IEEE*, 98(8), 1363–1366.

Poruschi, L., and Ambrey, C. L. (2019). “Energy justice, the built environment, and solar photovoltaic (PV) energy transitions in urban Australia: A dynamic panel data analysis.” *Energy Research and Social Science*, Elsevier, 48(October 2018), 22–32.

R Core Team. (2017). “R: A Language and Environment for Statistical Computing.” *R Foundation For Statistical Computing*.

Rafsanjani, H. N., Ahn, C., and Chen, J. (2017). “Analysis of Delay Interval and Energy-Load Variation for Non-Intrusively Extracting Occupant Energy-Use Information in Commercial Buildings.” *ASCE International Workshop on Computing in Civil Engineering 2017*, 191–197.

Ramachandra, B., Nawathe, P., Monroe, J., Han, K., Ham, Y., Vatsavai, R. R., Ramachandra, B., Vatsavai, R. R., and Nawathe, P. (2018). "Real-Time Energy Audit of Built Environments: Simultaneous Localization and Thermal Mapping." *Journal of Infrastructure Systems*, 24(3), 04018013.

Ratti, C., Raydan, D., and Steemers, K. (2003). "Building form and environmental performance: Archetypes, analysis and an arid climate." *Energy and Buildings*, 35(1), 49–59.

Reckien, D., Salvia, M., Heidrich, O., Church, J. M., Pietrapertosa, F., De Gregorio-Hurtado, S., D'Alonzo, V., Foley, A., Simoes, S. G., Krkoška Lorencová, E., Orru, H., Orru, K., Wejs, A., Flacke, J., Olazabal, M., Geneletti, D., Feliu, E., Vasilie, S., Nador, C., Krook-Riekkola, A., Matosović, M., Fokaides, P. A., Ioannou, B. I., Flamos, A., Spyridaki, N. A., Balzan, M. V., Fülöp, O., Paspaldzhiev, I., Grafakos, S., and Dawson, R. (2018). "How are cities planning to respond to climate change? Assessment of local climate plans from 885 cities in the EU-28." *Journal of Cleaner Production*, 191, 207–219.

Roth, J., and Jain, R. K. (2018). "Data-Driven, Multi-metric, and Time-Varying (DMT) Building Energy Benchmarking Using Smart Meter Data." *Workshop of the European Group for Intelligent Computing in Engineering*, 568–593.

Sanaieian, H., Tenpierik, M., Linden, K. Van Den, Mehdizadeh Seraj, F., and Mofidi Shemrani, S. M. (2014). "Review of the impact of urban block form on thermal performance, solar access and ventilation." *Renewable and Sustainable Energy Reviews*, 38, 551–560.

Sanguinetti, A., Dombrovski, K., and Sikand, S. (2018). "Information, timing, and display: A design-behavior framework for improving the effectiveness of eco-feedback." *Energy Research and Social Science*, 39(October 2017), 55–68.

Sarralde, J. J., Quinn, D. J., Wiesmann, D., and Steemers, K. (2015). "Solar energy and urban morphology: Scenarios for increasing the renewable energy potential of neighbourhoods in London." *Renewable Energy*, 73, 10–17.

Schelly, C. (2014). "Residential solar electricity adoption: What motivates, and what matters? A case study of early adopters." *Energy Research and Social Science*, Elsevier Ltd., 2, 183–191.

Van Der Schoor, T., and Scholtens, B. (2015). "Power to the people: Local community initiatives and the transition to sustainable energy." *Renewable and Sustainable*

Energy Reviews, Elsevier, 43, 666–675.

Schot, J., Kanger, L., and Verbong, G. (2016). "The roles of users in shaping transitions to new energy systems." *Nature Energy*, 1, 1–7.

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., and Griskevicius, V. (2007). "The Constructive, Destructive, and Reconstructive Power of Social Norms." *Psychological Science*, 18(5), 429.

"Seattle Energy Benchmarking." (2018). *Office of Sustainability & Environment*, <http://www.seattle.gov/energybenchmarkingmap>.

Seyfang, G., Haxeltine, A., Hargreaves, T., and Longhurst, N. (2010). *Energy and communities in transition - towards a new research agenda on agency and civil society in sustaianbility transitions*. *CSERGE working paper EDM 10-13*.

SF Environment. (2012). *Protecting Solar Access*.

Sharp, T. (1995). "Energy Benchmarking In Commercial Office Buildings." *ACEEE Summer Study on Energy Efficiency in Buildings*, 4, 321–329.

Shelton, T., Zook, M., and Wiig, A. (2015). "The 'actually existing smart city.'" *Cambridge Journal of Regions, Economy and Society*, 8(1), 13–25.

Shirer, M., and Da Rold, S. (2018). "IDC Forecasts Smart Cities Spending to Reach $158 Billion in 2022, with Singapore, Tokyo, and New York City Among Top Spenders." *International Data Corporation*, <https://www.idc.com/> (Sep. 9, 2018).

Shrestha, P. P., and Kulkarni, P. (2013). "Factors influencing energy consumption of energy star and non-energy star homes." *Journal of Management in Engineering*, 29(3), 269–278.

Siemens. (2020). "Dimensions of smart city development." *Siemens*, <https://new.siemens.com/global/en/company/topic-areas/smart-infrastructure/smart-cities.html>.

SierraClub. (2019). "100% Commitments in Cities, Counties, & States." *Sierra Club*, <https://www.sierraclub.org/ready-for-100/commitments> (Feb. 10, 2019).

Sigrin, B., and Mooney, M. (2018). *Rooftop Solar Technical Potential for Low-to-Moderate Income Households in the United States*. *National Renewable Energy Laboratory*.

Skjølsvold, T. M., Jørgensen, S., and Ryghaug, M. (2017). “Users, design and the role of feedback technologies in the Norwegian energy transition: An empirical study and some radical challenges.” *Energy Research and Social Science*, Elsevier Ltd, 25, 1–8.

Skjølsvold, T. M., and Lindkvist, C. (2015). “Ambivalence, designing users and user imaginaries in the European smart grid: Insights from an interdisciplinary demonstration project.” *Energy Research and Social Science,* Elsevier Ltd, 9, 43–50.

Solomon, S., Plattner, G. K., Knutti, R., and Friedlingstein, P. (2009). “Irreversible climate change due to carbon dioxide emissions.” *Proceedings of the National Academy of Sciences of the United States of America*, 106(6), 1704–1709.

Sorrentino, M., Bruno, M., Trifirò, A., and Rizzo, G. (2019). “An innovative energy efficiency metric for data analytics and diagnostics in telecommunication applications.” *Applied Energy*, Elsevier, 242(March), 1539–1548.

Sovacool, B. K. (2014). “What are we doing here? Analyzing fifteen years of energy scholarship and proposing a social science research agenda.” *Energy Research and Social Science*, Elsevier Ltd., 1, 1–29.

Squire, K., and Klopfer, E. (2007). *Augmented reality simulations on handheld computers*. *Journal of the Learning Sciences*.

Strengers, Y. (2014). “Smart Energy in Everyday Life: Are You Designing for Resource Man?” *Interactions*, 4, 24–31.

“Sun Number.” (2020). *Sun Number*, <https://www.sunnumber.com/> (Feb. 2, 2020).

Sunter, D. A., Castellanos, S., and Kammen, D. M. (2019). “Disparities in rooftop photovoltaics deployment in the United States by race and ethnicity.” *Nature Sustainability*, 2(1), 71–76.

Sustainable Energy Roadmap. (2016). *Solar PV Master Planning for Multi-Family Buildings*.

Sutherland, M. K., and Cook, M. E. (2017). “Data-driven smart cities: A closer look at organizational, technical & data complexities.” *ACM International Conference Proceeding Series*, Part F1282, 471–476.

Tu, K.-J. (2015). “Establishing the DEA energy management system for individual departments within universities.” *Facilities*, 33(11/12), 716–735.

U.S. EIA. (2020). *Annual Energy Outlook 2020. U.S. Energy Information Administration*.

United Nations. (2019). “Status of Ratification.” *United Nations Framework Convention on Climate Change*, <https://unfccc.int/process/the-paris-agreement/status-of-ratification>.

US Census. (2019). “Metropolitan and Micropolitan.” *US Census*, <https://www.census.gov/programs-surveys/metro-micro.html>.

Valkanova, N., Jorda, S., Tomitsch, M., and Vande Moere, A. (2013). “Reveal-it!: the impact of a social visualization projection on public awareness and discourse.” *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3461–3470.

Wang, S., Kim, A. A., and Johnson, E. M. (2017). “Understanding the deterministic and probabilistic business cases for occupant based plug load management strategies in commercial office buildings.” *Applied Energy*, 191, 398–413.

Wang, S., Yan, C., and Xiao, F. (2012). “Quantitative energy performance assessment methods for existing buildings.” *Energy and Buildings*, Elsevier B.V., 55, 873–888.

Wang, Y., Chen, Q., Hong, T., and Kang, C. (2018). “Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges.” *IEEE Transactions on Smart Grid*, 3053(June 2017), 1–24.

We Are Still In. (2020). “President Trump Wants Out - We Are Still In.” <https://www.wearestillin.com/>.

Weiss, M., Mattern, F., Graml, T., Staake, T., and Fleisch, E. (2009). “Handy feedback: Connecting smart meters with mobile phones.” *Paper presented at the 8th International Conference on Mobile and Ubiquitous Multimedia*.

Wood, G., and Newborough, M. (2007). “Energy-use information transfer for intelligent homes: Enabling energy conservation with central and local displays.” *Energy and Buildings*, 39(4), 495–503.

World Health Organization. (2016). “WHO Global Urban Ambient Air Pollution Database.” *World Health Organization*, <https://www.who.int/airpollution/data/cities-2016/en/>.

World Meteorological Organization. (2019). *WMO Greenhouse Gas Bulletin*. *World Meteorological Organization*.

Wu, H. K., Lee, S. W. Y., Chang, H. Y., and Liang, J. C. (2013). “Current status, opportunities and challenges of augmented reality in education.” *Computers and Education*, Elsevier Ltd, 62, 41–49.

Xu, X., Taylor, J. E., and Pisello, A. L. (2014). “Network synergy effect: Establishing a synergy between building network and peer network energy conservation effects.” *Energy and Buildings*, Elsevier B.V., 68, 312–320.

Xuchao, W., Priyadarsini, R., and Siew Eang, L. (2010). “Benchmarking energy use and greenhouse gas emissions in Singapore’s hotel industry.” *Energy Policy*, Elsevier, 38(8), 4520–4527.

Yarbrough, I., Sun, Q., Reeves, D. C., Hackman, K., Bennett, R., and Henshel, D. S. (2015). “Visualizing building energy demand for building peak energy analysis.” *Energy and Buildings*, Elsevier B.V., 91, 10–15.

Yu, J., Wang, Z., Majumdar, A., and Rajagopal, R. (2018). “DeepSolar: A Machine Learning Framework to Efficiently Construct Solar Deployment Database in the United States.” *Joule*, Elsevier Inc., 2(12), 2605–2617.

Zanella, a, Bui, N., Castellani, a, Vangelista, L., and Zorzi, M. (2014). “Internet of Things for Smart Cities.” *IEEE Internet of Things Journal*, 1(1), 22–32.

Zhang, J., Xu, L., Shabunko, V., Tay, S. E. R., Sun, H., Lau, S. S. Y., and Reindl, T. (2019). “Impact of urban block typology on building solar potential and energy use efficiency in tropical high-density city.” *Applied Energy*, 240(February), 513–533.

Zhang, Y., O’Neill, Z., Dong, B., and Augenbroe, G. (2015). “Comparisons of inverse modeling approaches for predicting building energy performance.” *Building and Environment*, Elsevier Ltd, 86, 177–190.

Zhao, D., McCoy, A. P., Du, J., Agee, P., and Lu, Y. (2017). “Interaction effects of building technology and resident behavior on energy consumption in residential buildings.” *Energy and Buildings*, Elsevier B.V., 134, 223–233.

Zhou, K., Fu, C., and Yang, S. (2016). “Big data driven smart energy management: From big data to big insights.” *Renewable and Sustainable Energy Reviews*, Elsevier Ltd, 56(2016), 215–225.

Zhou, K., and Yang, S. (2016). “Understanding household energy consumption behavior: The contribution of energy big data analytics.” *Renewable and Sustainable Energy Reviews*, Elsevier, 56, 810–819.

Zhu, R., You, L., Santi, P., Wong, M. S., and Ratti, C. (2019). “Solar accessibility in developing cities: A case study in Kowloon East, Hong Kong.” *Sustainable Cities and Society*, Elsevier, 51(July), 101738.

Zuiderwijk, M., Janssen, R., and Matheus, A. (2015). “Big and Open Linked Data (BOLD) to Create Smart Cities and Citizens: Insights from Smart Energy and Mobility Cases.” *Electronic Government*, Springer, Cham, 79–90.

Zullo, L., Resources, N., Council, D., Antonoff, J., and Amberli, Y. (2016). “Putting Building Energy Benchmarking Data into Action.” *ACEEE Summer Study on Energy Efficiency in Buildings*, 1–12.

Zuo, J., Read, B., Pullen, S., and Shi, Q. (2013). “Carbon-neutral commercial building development.” *Journal of Management in Engineering*, 29(1), 95–102.